PHENOTYPIC DIVERSITY WITHIN TWO TOXIC DINOFLAGELLATE GENERA: ENVIRONMENTAL AND TRANSCRIPTOMIC STUDIES OF SPECIES DIVERSITY IN *ALEXANDRIUM* AND *GAMBIERDISCUS*

by

Kathleen Johnson Pitz

B. A., University of California Berkeley (2009)

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

September 2016

Author……....
Joint Program in Applied Ocean Sciences & Engineering
Massachusetts Institute of Technology
& Woods Hole Oceanographic Institution
August 16th, 2016

Certified by……
Donald M. Anderson
Senior Scientist
Woods Hole Oceanographic Institution
Thesis Supervisor

Accepted by……
Ann M. Tarrant
Chair, Joint Committee for Biological Oceanography
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution

PHENOTYPIC DIVERSITY WITHIN TWO TOXIC DINOFLAGELLATE GENERA: ENVIRONMENTAL AND TRANSCRIPTOMIC STUDIES OF SPECIES DIVERSITY IN *ALEXANDRIUM* AND *GAMBIERDISCUS*

by
Kathleen Johnson Pitz

Submitted to the MIT-WHOI Joint Program in Oceanography and Applied Ocean Science and Engineering on August 16th, 2016, in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Oceanography

ABSTRACT

Dinoflagellates are a diverse group of single-celled eukaryotic phytoplankton that are important for their unique genetics and molecular biology, the multitude of ecological roles they play, and the ability of multiple species to produce toxins that affect human and ecosystems health. Two dinoflagellate genera, *Alexandrium* and *Gambierdiscus* each contain species that can cause human poisoning syndromes, although the methods of toxin transfer, accumulation, and exposure are very different. *Gambierdiscus* is a benthic organism that produces lipophilic ciguatoxins that can bioaccumulate in coral reef fish and cause ciguatera fish poisoning (CFP) in human consumers. *Alexandrium* is a planktonic species that produces saxitoxins that can directly accumulate in shellfish and cause paralytic shellfish poisoning (PSP) in humans. Both genera contain multiple species that vary dramatically in toxicity and physiology. Through transcriptomic analysis, this thesis describes the genetic diversity present across dinoflagellates that produce saxitoxin, elaborating on differences in their complement of genes within the saxitoxin biosynthesis pathway. This study demonstrated retention and expression of some of these saxitoxin genes by non-toxic species within *Alexandrium,* as well as in *Gambierdiscus*, which does not produce saxitoxins. Furthermore it confirmed the presence of certain transcripts only in toxin-producing species. This thesis then developed novel fluorescence in situ hybridization (FISH) probes that can be used to identify and enumerate six *Gambierdiscus* species, thereby enabling the community composition of *Gambierdiscus* to be examined in a quantifiable manner. The probes were tested in the laboratory on cultures, and then successfully applied to field samples from Florida Keys and Hawai'i. *Gambierdiscus* species are diverse in both their toxicity and optimal temperature ranges for growth. Analysis of *Gambierdiscus* community composition in an area of variable temperature allowed the characterization of species shifts that were driven both by a seasonal increase in mean seawater temperatures and spatial variability of temperature experienced between tidal pools. Overall this thesis advances the knowledge of dinoflagellate genetics and ecology, aids in the characterization of species harmful to public health, and provides tools and approaches to help monitor and manage harmful effects from these species, including some that are projected to increase with climate change.

Thesis Supervisor: Donald M. Anderson
Woods Hole Oceanographic Institution

For my parents, Penny and Bill

ACKNOWLEDGMENTS

First I would like to thank my funding sources that have supported my research throughout my graduate education. I was privileged to receive support from the United States Food and Drug Administration (USFDA) (F223201000060C); the National Oceanic and Atmospheric Administration National Ocean Service (NOAA NOS) (CiguaHAB; NA11NOS4780060 and NA11NOS4780028); Woods Hole Center for Oceans and Human Health (COHH) National Science Foundation (NSF) Grant OCE-1314642 and National Institute of Environmental Health Sciences (NIEHS) 1P01-ES021923-01. I also received funding through the WHOI Ocean Ventures Fund and WHOI Coastal Institute Research Proposal Award. I was able to present this research at multiple conferences thanks to funds from the Academic Programs Office at WHOI, the MIT Student Assistance Fund, and the NOAA Center for Sponsored Coastal Ocean Research (CSCOR) handled through the U.S. National Office for Harmful Algal Blooms.

Thank you to the Anderson lab for welcoming and supporting me from the moment when I first got off the bus in Woods Hole to turning in my thesis. The last six years have been challenging but rewarding and I owe my success to your support both inside and outside of the lab. Don thank you for guiding my progress and research questions while still allowing me a great deal of intellectual freedom. I appreciate that from the beginning of graduate school you encouraged me to not only conduct my own experiments in the lab but offered me opportunities to visit field sites so I could gain a better appreciation of the ecology of the system I was working in. I would also like to thank my thesis committee for many productive conversations and help improving my thesis: Sonya Dyhrman, Edward DeLong, Rebecca Gast, and Robert Dickey.

I owe thanks to all of the Anderson lab over the years but I would like to especially thank Mindy for first introducing me to *Gambierdiscus* and giving me so much support both in the lab and the field since; Mike for encouraging me while I tackled transcriptomic work, and being patient while I hit multiple roadblocks; Dave for handling most of my day-to-day questions about experimental procedures and for teaching me how to keep dinoflagellates alive and happy; and Judy for not only handling administrative headaches but for also giving out great gardening and cooking advice. Thank you also to everyone who had a hand in helping me process field samples (especially Taylor), this research wouldn't have been possible without you. Finally, special thanks are due to Alexis for being a great partner in crime and officemate, and tolerating my eccentricities for five years. I am also grateful for all the help I received outside of the Anderson lab ranging from field support and lab space in Hawai'i (thanks especially to Jason Adolf, Louise Economy, and Ashley Brandt); preserved field samples and *Gambierdiscus* research conversations (Mike Parsons, Tyler Smith, and Alison Robertson); and help with statistical analysis (Andy Beet and Andy Solow).

I would like to thank my friends both inside and outside of Woods Hole. Thank you for the hikes, swim practices, bonfires and board games. You've made me a stronger scientist and person. I've grown so much with you. Thank you to my biology cohort; I couldn't imagine better people to have shared this experience with. I look forward to many more laughs and adventures in the future. Special thanks to Josh for love and support through this process; I couldn't have made it through the end without you. Finally, thank you to my family who encouraged my love of nature from an early age and gave me many opportunities to see all the fish, mountains, tidepools, and wildlife I could want.

CONTENTS

Chapter 1

Introduction

Background and Motivation

Dinoflagellates are a diverse group of phytoplankton both in their genetic cell biology and physiology. All have incredibly large and unusually structured genomes (Keeling & Slamovits, 2005; Wisecaver & Hackett, 2011). They fill important niches in the environment – some are phototrophs, some are heterotrophs, and a growing fraction are recognized as intermediates - mixotrophic organisms (Jeong et al., 2010). Dinoflagellates inhabit a variety of environments from the open ocean to the benthos, or ocean bottom. Multiple dinoflagellate species are well known as significant hazards to human health through their production of toxins that can accumulate in seafood. The cell proliferations that lead to these poisoning syndromes are termed harmful algal blooms or HABs. HABs occur in both freshwater and marine environments. Some HABs contaminate seafood through direct accumulation (e.g. in shellfish) or via bioaccumulation, bioconversion, and biomagnification through the marine food chain. Since some dinoflagellates produce toxins, their study is important not only for the unique nature of their basic cellular biology and important role in the environment, but also for human health applications and the effects their toxin production has on co-occurring organisms and ecosystems.

Two dinoflagellate genera, *Alexandrium* and *Gambierdiscus*, each contain species that can cause poisoning syndromes although the methods of toxin accumulation and the symptomology of the poisoning syndromes are different. *Gambierdiscus* is a benthic organism that produces lipophilic ciguatoxins that can bioaccumulate in coral reef fish and cause ciguatera fish poisoning (CFP) in human consumers. *Alexandrium* is a planktonic species that produces saxitoxins that can directly accumulate in shellfish and cause paralytic shellfish poisoning (PSP) in humans. Both genera contain multiple species that vary dramatically in toxicity and physiology.

This thesis first uses the saxitoxin gene cluster in *Alexandrium* to illustrate how such closely related dinoflagellate species become divergent in toxicity, which specific components of this cluster are necessary for toxicity, and the extent to which this gene cluster illustrates the diversification processes of dinoflagellate genomics. It also shows how this diversification can result in differences in physiology or phenotypes. This thesis then focuses on the genus *Gambierdiscus* to answer questions of how diverse species within the same genus occupy environmental niches on a small geographic scale. Similar to *Alexandrium*, species of *Gambierdiscus* are divergent in toxin production and also have different tolerances to abiotic factors such as temperature. By examining *Gambierdiscus* community composition in Wai'Ōpae Tide Pools, Hawai'i, a region of shifting temperatures, this thesis determined how *Gambierdiscus* species composition could shift through time as processes such as climate change alter environmental conditions. A critical tool in the analysis of the community diversity of *Gambierdiscus* in Wai'Ōpae was a set of fluorescent in situ hybridization (FISH) probes that were developed (Chapter 3), allowing identification and quantification of six species within mixed field assemblages.

Motivating Questions:

1. What is the genetic basis for saxitoxin production in dinoflagellates?
2. What components of the saxitoxin gene cluster are necessary for toxicity?
3. How does the saxitoxin gene cluster illustrate the diversification processes of dinoflagellate gene clusters?
4. What is the species diversity of natural populations of *Gambierdiscus*?
5. How does the species assemblage and abundance of *Gambierdiscus* shift through time?
6. How do changing environmental parameters, such as temperature, affect *Gambierdiscus* population structure?

Paralytic Shellfish Poisoning

Paralytic shellfish poisoning (PSP) is a harmful algal bloom (HAB) threat whereby contamination of seafood occurs through direct accumulation of phycotoxins, saxitoxin and its derivatives, named paralytic shellfish toxins (PSTs) (Cembella, 1998). These are a family of neurotoxins produced by both prokaryotic cyanobacteria and eukaryotic dinoflagellates. Saxitoxin and its congeners act by binding and inhibiting voltage-gated sodium (Na+) ion channels (Kao & Nishiyama, 1965). These Nav channels are present in mammalian neurons (Catterall, 2012). In humans, inhibition of these channels by PSTs can lead to paralysis and death (Llewellyn, 2006). In the United States, toxicity in shellfish is heavily monitored and thus rarely leads to human fatalities, but it has vast economic effects due to shellfish harvesting closures and loss of tourism and recreational water use (Anderson et al., 2012). Within wild animal populations, PSP remains responsible for mortality events as well as ecosystem-level restructuring as it influences multiple trophic levels such as marine mammals, birds, fish, and invertebrates (Jensen et al., 2015; Shumway, Allen, & Dee Boersma, 2003; Zimmer & Ferrer, 2007).

Several dinoflagellates within the *Alexandrium* genus produce saxitoxins. In total the dinoflagellate genus *Alexandrium* consists of over 30 defined species, half of which have been shown to have toxic effects (Anderson et al., 2012). The *A. tamarense* species complex comprises five of these species and contains both toxic and non-toxic species in a closely related cluster. Of the five species in the *Alexandrium tamarense* species complex, three are shown to be toxic (*A. fundyense*, *A. pacificum,* and *A. australiense*) and two are non-toxic or produce toxin below the limits of detection (*A. mediterraneum* and *A. tamarense*) (John et al., 2014). The *A. tamarense* species complex presents a unique opportunity to investigate the saxitoxin biosynthesis pathway since its members contain both toxic and non-toxic species in a closely related cluster. In Chapter 2, by examining relationships among the *Alexandrium tamarense* species complex and their complement of saxitoxin genes, differences are established between the toxic and non-toxic species, illustrating genetic changes that may be important for toxin biosynthesis and further elucidating the complexity of the saxitoxin gene cluster in dinoflagellate transcriptomes.

A. tamarense species complex

The *A. tamarense* species complex is not only responsible for the majority of toxic blooms caused by *Alexandrium* species (Emily L. Lilly, Halanych, & Anderson, 2007) but due to its large geographical range, is also the primary marine dinoflagellate species responsible for PSP risk (Anderson et al., 2012), which in turn is the most widespread of all HAB human poisoning

syndromes. The *tamarense* complex, formally known as the tamarensis complex, was created when examination of morphological characteristics failed to consistently differentiate between *Alexandrium tamarense*, *catenella*, and *fundyense* species (Balech 1985). Five ribotypes were defined and preliminarily named after their geographical origin (Scholin, Herzog, Sogin, & Anderson, 1994). Recent research describing a greater number of globally distributed sequences has confirmed the clustering, renamed the ribotypes numerically as Groups I-V, and initially shown that these ribotypes exclusively include either toxic or non-toxic member species (Emily L. Lilly et al., 2007). These ribotypes differ from each other by 13-18% in their ITS1/5.8S/ITS2 sequences (Orr, Stüken, Rundberget, Eikrem, & Jakobsen, 2011) and offspring from the mating of two of the ribotypes are inviable (Brosnahan et al., 2010). These ribotypes were then characterized as separate species and given the names *A. fundyense* (Group I), *A. mediterraneum* (Group II), *A. tamarense* (Group III), *A. pacificum* (Group IV), and *A. australiense* (Group V) (John et al., 2014). These ribotypes are thus today considered distinct species that have their own physiological requirements and differences in toxicity.

Saxitoxin biosynthesis pathway

The saxitoxin biosynthesis pathway in cyanobacteria and dinoflagellates has been under a high level of scrutiny due to its economic importance and the increasing availability of sequence data from saxitoxin-producing organisms. Furthermore the chemical structure of saxitoxin is interesting in itself: it consists of a tricyclic frame with two guanidine groups and more heteroatoms than carbon atoms (Tsuchiya et al., 2016).

The mechanism of saxitoxin biosynthesis was first discovered through biochemical feeding experiments with radiolabeled amino acids and $^{13}CO_2$ in *Alexandrium* and *Aphanizomenon flos aquae* (Shimizu et al., 1993; Shimizu, Norte, Hori, Genenah, & Kobayashi, 1984). With the advent of sequencing technology, the elaboration of the first saxitoxin biosynthetic gene cluster in cyanobacteria took place (Kellmann et al., 2008), and quickly following were discoveries of homologous sequences from diverse saxitoxin-producing dinoflagellate species, including several *A. tamarense* complex species, *Gymnodinium catenatum*, and *Pyrodinium bahamense* (Hackett et al., 2012; Murray, Diwan, Orr, Kohli, & John, 2015; Orr, Stüken, Murray, & Jakobsen, 2013a, 2013b; Stüken et al., 2011).

In cyanobacteria, the saxitoxin gene pathway has been defined in the species *Cylindrospermopsis raciborskii*, *Anabaena circinalis*, *Aphanizomenon sp.*, *Raphidiopsis brookii*, and *Lyngbya wollei* (Kellmann et al., 2008; Mihali, Carmichael, & Neilan, 2011; Mihali, Kellmann, & Neilan, 2009; Moustafa et al., 2009; Stucken et al., 2010). In *Cylindrospermopsis raciborskii* the saxitoxin biosynthesis gene cluster is 35kbp long with 31 open reading frames (Kellmann, Stüken, Orr, Svendsen, & Jakobsen, 2010). Twenty-six proteins have been found to complete the cyanobacterial saxitoxin pathway, named sxtA-Z (Kellmann et al., 2008). Due to the similarity in their saxitoxin biosynthetic pathways and the stereochemistry of their toxin products, it is likely that cyanobacteria and dinoflagellates both use homologous enzymes to synthesize saxitoxin (Kellmann et al., 2010). This similarity allows cyanobacterial saxitoxin genes to be used in discovery of the dinoflagellate saxitoxin biosynthesis pathway. Among the five cyanobacterial strains studied there is a set of 14 core proteins that are present in every toxic strain (sxtA-sxtI, sxtP-sxtR, sxtS, and sxtU) (Murray, Mihali, & Neilan, 2011) and 8 proteins which are considered

directly involved in saxitoxin biosynthesis (sxtA, sxtB, sxtD, sxtG, sxtH, sxtT, sxtI, sxtS, and sxtU) (Kellmann et al., 2008).

Two genes have been identified as important to saxitoxin biosynthesis in dinoflagellates: sxtA and sxtG. They are the first two genes to act in the saxitoxin biosynthesis pathway and it is hypothesized their presence or absence can signify if a dinoflagellate species has the ability to produce saxitoxin (Murray, Wiese, et al., 2011; Orr et al., 2013b; Wiese, Murray, Alvin, & Neilan, 2014). In a study by Stüken et al. (2011) the nucleotide and protein sequences of two isoforms of sxtA were described in an *Alexandrium fundyense* strain. An additional study by Hackett et al (2012) found five protein sequences related to sxtA and two protein sequences related to sxtG in an *Alexandrium pacificum* strain. The same study found dinoflagellate sequences which corresponded to thirteen proteins of the saxitoxin gene cluster, including eight of the proteins directly involved in the biosynthesis of saxitoxin in the same strain (sxtA, sxtB, sxtD, sxtF/M, sxtG, sxtH/T, sxtI, sxtL, sxtN, sxtP, sxtS, sxtU, and sxtX) (Hackett et al., 2012).

Importantly, the sxtA gene in cyanobacteria consists of four domains, sxtA1-4, including a SAM-dependent methyltransferase, GCN5- related N-acetyltransferase, acyl carrier protein and class II aminotransferase (Kellmann et al., 2008). Only the last domain, sxtA4, has been only found in toxic and not non-toxic dinoflagellate species. Several dinoflagellate homologs to sxtA have been identified containing varying sxtA domains: from Stüken et al., 2011, of the two sxtA isoforms the first contains domains sxtA1-3 and the second longer transcript contains domains sxtA1-4; from Hackett et al., 2012, of the two additional putative sxtA homologs one contains domains similar to sxtA1-3 and the other contains domains similar to sxtA2-3. In summation only one dinoflagellate homolog has been discovered to contain the sxtA4 domain. This domain has been hypothesized to be essential to the dinoflagellate production of saxitoxin and has been reported across saxitoxin-producing strains of dinoflagellates outside the *A. tamarense* species complex as well (Murray et al., 2015).

There have been some conflicting reports on the presence or absence of sxtA4 and sxtG in non-toxic *A. tamarense*. PCR studies of the presence of this domain in *A. fundyense*, *A. tamarense*, *A. pacificum*, and *A. australiense* have found it present in gDNA across all strains and also in cDNA from *A. fundyense*, *A. pacificum*, and *A. tamarense* (Stüken et al., 2011). However this finding of sxtA4 in non-toxic *A. tamarense* has not been repeated. In a study by Murray et al., 2011, a qPCR assay of Group III gDNA did not amplify sxtA4 (Murray et al., 2011) and Murray et al, 2015 did not find this domain in *A. tamarense* strains (Murray et al., 2015). Additionally, sxtG, the second gene to act in saxitoxin biosynthesis, was amplified in *A. tamarense* cDNA and gDNA by PCR yet a qPCR assay on gDNA showed no reaction or was unspecific (Orr et al., 2013). Chapter 2 presents additional evidence that sxtA4 and sxtG are either not expressed or expressed at a very low levels within *A. tamarense* and *A. mediterraneum*, the two non-toxic members of the *A. tamarense* species complex.

Assays have been developed to detect the domain sxtA4 in the environment (Murray, Wiese, et al., 2011; Stüken et al., 2013) in order to determine if there are potential saxitoxin-producers present as a method of monitoring for risk of PSP. However, recent biochemical research has discovered a shunt pathway which is used in saxitoxin biosynthesis to decrease PST production in the cyanobacterium *A. circinalis*, exporting an intermediate product into the extracellular environment (Tsuchiya et al., 2016). This shunt pathway occurs after the action of sxtA, sxtG

and sxtB, and directly previous to the proposed action of sxtD and sxtS according to the pathway proposed by Kellmann et al (Kellmann et al., 2008). The intermediate shunt product has also been found in *Alexandrium*, and if this pathway is found to operate in dinoflagellates, it would further complicate the ability of a qPCR assay for sxtA to predict the presence of PSTs in the environment. Such an assay would still discriminate between species that have the ability to produces PSTs but not between levels of toxicity of toxin producing cells.

A hypothesis explaining how organisms from two different kingdoms of life, eukaryotic marine dinoflagellates and prokaryotic freshwater cyanobacteria, came to both produce saxitoxins is that a horizontal gene transfer (HGT) event transferred cyanobacterial saxitoxin genes between a common dinoflagellate ancestor of both *Alexandrium* and *Pyrodinium,* and a cyanobacterium (Murray et al., 2015; Orr et al., 2013a). Furthermore, a second HGT event could have occurred between an *Alexandrium tamarense* species and *Gymnodinium catenatum* to explain its paraphyletic production of the toxin (Orr et al., 2013) or alternatively the first HGT event occurred long enough in the past to have involved the ancestor of *A. tamarense* and *G. catenatum* and cyanobacterial genes were later lost from other non-toxic lineages (Murray et al., 2015). This would lead us to expect remnant saxitoxin biosynthesis genes to potentially be present in at least species sharing a common ancestor with *Alexandrium* and *Pyrodinium*. SxtA genes have been described in a common ancestor with *Gymnodinium catenatum* and *Alexandrium*, *Scrippsiella trochoidea* (Cooper, Sinclair, & Wawrik, 2016). Fourteen out of twenty-six saxitoxin biosynthesis genes in the sxtA-Z pathway were also found. This presents additional evidence that there may have been a HGT event that occurred further back in history.

Horizontal gene transfer events are common in dinoflagellates: a study found that *A. tamarense* had the highest numbers of genes resulting from HGT events including aveolates, stramenopiles, a haptophyte, a cryptophyte, and prasinophytes (Wisecaver, Brosnahan, & Hackett, 2013). Furthermore, dinoflagellates have been shown to 'recycle' genetic material to regulate their plastid photosynthetic machinery (Méheust, Zelzion, Bhattacharya, Lopez, & Bapteste, 2016). In the case of the saxitoxin biosynthesis pathway, it's possible that genes from a HGT event between a cyanobacterium and a dinoflagellate are now being used not only to produce saxitoxin and its derivatives, but also may be used in other pathways.

In Chapter 2 I looked for components of the saxitoxin biosynthesis pathway that illustrate which sequences may be most important for toxin production using a natural system of species that are related, yet divergent in toxin production. I examined the presence of saxitoxin genes across the five species within the *Alexandrium tamarense* species complex (*A. fundyense*, *A. mediterraneum*, *A. tamarense*, *A. australiense*, and *A. pacificum)*, and other toxic and non-toxic species within the genus *Alexandrium* (*A. ostenfeldii*, *A. affine*, *A. monilatum, A. minutum, and A. andersonii),* as well as saxitoxin-producing species *Pyrodinium bahamense* and *Gymnodinium catenatum* and five *Gambierdiscus* species: *G. australes, G. belizeanus, G. caribaeus, G. pacificus, G. silvae and Gambierdiscus sp. type 4*. Since *Gambierdiscus* shares a common ancestor with *Pyrodinium* and *Alexandrium* but does not produce saxitoxins I expected to find remnants of saxitoxin genes but not a functioning gene cluster. I showed that genes within the saxitoxin biosynthesis pathway have been retained and expressed within non-toxic dinoflagellates, including *Gambierdiscus*. Furthermore, that there has been a high diversification of transcripts, especially within sxtA, which illustrates how dinoflagellates can incorporate and diversify new genetic material to fill other cellular functions. Additionally, I found clusters of transcripts that are only present in toxic species, indicating their importance to

saxitoxin production and further elucidating the genetic basis for this toxin production within dinoflagellates.

Ciguatera Fish Poisoning

Ciguatera fish poisoning (CFP), or 'ciguatera', is the most prominent example of a HAB threat where bioaccumulation, bioconversion, and biomagnification of a phycotoxin through the marine food chain causes contamination of seafood. CFP is a syndrome caused by the bioaccumulation of lipophilic ciguatoxins in coral reef fish and subsequent consumption by humans (Scheuer, Takahashi, Tsutsumi, & Yoshida, 1967). These phycotoxins are produced by *Gambierdiscus*, a tropical epiphytic dinoflagellate genus that lives on many varieties of macroalgae but also may occur on dead corals and sand (Yasumoto, Nakajima, Bagnis, & Adachi, 1977). Globally, tens of thousands of individuals are likely afflicted with ciguatera on an annual basis, with up to 10% of the local population on some islands in endemic areas becoming ill (Lora E. Fleming, Baden, Bean, Weisman, & Blythe, 1998). CFP is also heavily underreported, as in endemic regions it is believed that there is no relief from medical treatment, and in non-endemic regions there is a lack of diagnostic recognition of the disease (Dickey & Plakas, 2010).

Since the initial discovery of the structure of the main Pacific Ocean ciguatoxin (P-CTX) and a precursor ciguatoxin produced by *Gambierdiscus* (Murata, Legrand, Ishibashi, Fukui, & Yasumoto, 1990), additional ciguatoxins have been described from the Caribbean and Indian Ocean (Hamilton, Hurbungs, Jones, & Lewis, 2002; Lewis, Vernoux, & Brereton, 1998). Ciguatoxins function as sodium channel agonists and consist of many congeners. They are lipid-soluble polyether compounds that can accumulate in fish flesh and are hypothesized to also accumulate within humans, as symptoms from repeated poisonings can be more severe (Bagnis, Kuberski, & Laugier, 1979). CFP can cause a diverse range of symptoms from gastrointestinal pain, vomiting, parasthesias and neurological symptoms such as the reversal of hot and cold sensations (dysesthesias) or chronic fatigue that can persist for months to years in extreme cases (Dickey & Plakas, 2010). These symptoms can sporadically reoccur with the consumption of fish, alcohol, or with an increase in exercise. Furthermore, some symptoms are more common of certain geographic areas, as CFP in the Pacific and Indian Oceans is reputed to have more neurological symptoms. This difference in presentation of the syndrome is likely due to the diversity of ciguatoxins present in the environment, the diversity of toxic fish consumed, and potentially the different geographic ranges of *Gambierdiscus* species. The different paths these toxins traverse the marine food web may cause a varying array of symptoms as ciguatoxins undergo bioconversion in fish, allowing different suites of ciguatoxins to accumulate in predatory verses herbivorous fish for example (Lewis & Holmes, 1993). There are even human modes of transmission as CFP has been documented to be sexually transmitted (Lange, Lipkin, & Yang, 2008).

Ciguatera is considered endemic in much of the tropical world. On many islands, reef fish represent the cheapest and healthiest protein source for local populations. This protein source is threatened by ciguatera. Historically, an increase in ciguatoxic fish may have driven the extensive Polynesian migrations between A.D. 1000-1450 as climate conditions favorable to ciguatera became more frequent (Rongo, Bush, & Van Woesik, 2009). Even small amounts of toxin, consumed over long periods of time, may eventually reach a threshold leading to symptoms of poisoning. CFP toxins are tasteless, odorless, and heat stable, therefore cooking a fish will not make it lose its toxicity (Baden, Fleming, & Bean, 1995). Among local island populations, some common methods of testing for toxicity include feeding a piece of fish to a cat and watching for

symptoms of poisoning, or seeing if ants or flies are attracted to the fish (Park, 1994). Ultimately, local populations rely on fishermen to know which reefs are toxic or not. Through a process of elimination, local knowledge has accumulated over which areas are generally safe to fish from and which have resulted in poisonings. Unfortunately with shifting environmental conditions due to global warming, many of these fishing histories may lose their accuracy as *Gambierdiscus* shifts in abundance due to temperature change, emphasizing the need to understand how increasing seawater temperatures and concomitant changes to reef communities influence *Gambierdiscus* population structure and physiology (Tester et al., 2010; Tosteson, 2004).

CFP outbreaks have long been associated with ecological disturbances to reef environments, likely due to proliferation of *Gambierdiscus'* preferred substrates of algae and dead coral (Kohlerl & Kohler, 1992). After nuclear testing in the Pacific, many islands where the coral reefs were damaged then suffered outbreaks of ciguatera (Ruff, 1989). This presents an added concern that as the health of global coral reefs suffer due to damage from overfishing, pollution, and bleaching events, the threat of CFP will increase concurrently. Due to its responsiveness to environmental disturbance, CFP has been described as a sensitive indicator of changing conditions in tropical marine ecosystems (Hales, Weinstein, & Woodward, 1999). In addition, its affect on human fishing behavior is also able to transform reef systems. Fishing of desirable, usually large, fish from reef ecosystems has a drastic effect across all trophic levels (Jennings & Kaiser, 1998). On reefs where fish are toxic, relaxation of fishing pressure can occur, effectively protecting reefs from overfishing. In other cases, where toxic fish exist but select species are still considered safe for consumption, a change in fishing pressure will also affect the reef. More ecological studies are needed to determine the site-specific effects of changing levels of toxicity on a reef, as an increase in herbivorous fish would presumably decrease macroalgal loads, decreasing *Gambierdiscus* habitat and therefore reef toxicity, while producing a cyclic effect back to high fishing pressures.

A critical aspect of calculating risk of ciguatera fish poisoning on the reef is *Gambierdiscus* community composition. Several studies have predicted that as global sea surface temperatures increase, *Gambierdiscus* will expand its geographic range, increasing the number of people at risk of ciguatera fish poisoning (Dickey & Plakas, 2010; P. A. Tester, Feldman, Nau, Kibler, & Wayne Litaker, 2010; Villareal et al., 2007). *Gambierdiscus* species have been shown to have different temperature ranges of growth as well as different ranges in toxicity (Holland et al., 2013; Kibler, Litaker, Holland, Vandersea, & Tester, 2012; Xu, Richlen, Liefer, Robertson, & Kulis, 2016). Therefore as seawater temperatures shift, *Gambierdiscus* species ranges will change, affecting not only total abundance of *Gambierdiscus* as a genus but also *Gambierdiscus* community composition. However, *Gambierdiscus* species are almost indistinguishable under the light microscope. Until recently, little was known of the genetic diversity present within the *Gambierdiscus* genus. The first molecular characterizations of *Gambierdiscus* species opened up opportunities for species enumeration (Litaker et al., 2009; Richlen, Morton, Barber, & Lobel, 2008). Current techniques to identify cells to the species level include sequencing of large subunit ribosomal RNA (LSU rRNA) gene or morphological analysis using scanning electron microscopy, both of which are impractical to apply on a large scale. Another technique using semi-quantitative qPCR has also been developed for large scale analysis however, due to varying levels of rDNA within *Gambierdiscus* cells, this assay does not have the same quantitative strength of microscope counts (Vandersea et al., 2012). In Chapter 3 and 4, I show that the development of species-specific fluorescent *in situ* hybridization (FISH) probes allow for analysis of field samples to determine *Gambierdiscus* community composition while preserving the accuracy of counting cells under the

light microscope. This differentiation of *Gambierdiscus* species allows the elaboration of community dynamics and shifting patterns of community toxicity and therefore elucidates the changing risk of ciguatera fish poisoning as well as providing a needed tool for monitoring of CFP risk.

Difficulty in Defining Dinoflagellate Species

Traditionally phytoplankton have been defined by morphological characteristics, such as plate morphology in the case of dinoflagellates. Recently these 'morphotypes' have been redefined using sequencing technology. Commonly this has resulted in a finer ability to discriminate between species, but has increased the difficulties involved in traditional monitoring. An important question to consider is at what level it becomes important to differentiate between organisms. In the case of dinoflagellates, this thesis shows significant genetic and physiological differences between closely related species of the same genus.

Our knowledge of eukaryotic microbial diversity in the ocean and its effects on ecosystems is expanding at an incredibly fast rate, partly in accordance with our ability to differentiate new levels of species diversity in the field. For example, simplistic characterization of genera of species, such as *Gambierdiscus*, as all one species masked the complexity of ranging toxicities and habitat preferences contained within its members. When evaluating the potential effects of climate change, or any habitat shift, our novel deeper understanding of genetic variability of these genera enables a more nuanced prediction. Essential to calculating risk of CFP is not only to consider the environmental changes that may increase or decrease *Gambierdiscus* abundance as a whole genus, but changes that may favor one species over others.

Novel Methods

Transcriptomic Analysis

Due to their large and chromosomally rich genomes, it remains largely impractical to sequence dinoflagellate genomes. Deep transcriptome sequencing circumvents this problem by only characterizing transcribed genes. By sequencing the transcriptomes of both *Gambierdiscus* and *Alexandrium,* the genetic basis for physiological differences between species within each genus can begin to be elaborated. Chapter 2 utilizes developing technologies available to analyze transcriptome data in order to capitalize on a new opportunity for genetic analysis of dinoflagellates.

Benefits of Transcriptomics compared with PCR

One strength of transcriptomic studies over PCR is the ability to pick up longer transcripts that can enable us to compare which domains are linked, such as in the case of sxtA in Chapter 2. In this chapter the diversity of sequences that contain domains similar to those of sxtA is greatly expanded. Furthermore, solely looking at PCR for a single domain region limits the analysis by the specificity of primers. In Chapter 2 through BLAST sequence similarity searches, I am able to find sequences that are more divergent and give insight into the evolutionary history of the saxitoxin biosynthesis pathway.

One limitation of transcriptomic work is the dependence on the quality of the transcriptomic data set. PCR is able to find rare sequences due to its specificity and amplification, but since transcriptomic analysis largely relies on the sequencing of total RNA, it's more likely a specific transcript will be missed if it is expressed at low levels. In Chapter 2 there is a range of transcriptome quality and completeness (as measured by *BUSCO* metrics (Simão, Waterhouse, Ioannidis, Kriventseva, & Zdobnov, 2015)). Therefore although some genes may be absent from a certain strain it does not necessarily mean that this is representative of the actual presence or absence of the transcript and transcriptome's quality is taken into account for this analysis. However, every additional sequencing project adds power to the collective understanding of an organism's gene content, and some of these transcriptomes are a first step towards a better understanding of an organism for which there is very little sequence data.

There has been limited success looking for differential transcription between the same dinoflagellate strains grown under different conditions and until recently, dinoflagellate transcriptomic regulation has remained difficult to observe. Instead, a hypothesis that strains are "hardwired" to produce a certain level of gene transcripts was put forward by the Palumbi lab at Stanford to explain coral's choice of dinoflagellate symbiont species (Barshis, Ladner, Oliver, & Palumbi, 2014). In this study, I propose that transcriptomics, while not letting us see the final protein impact, still allows us to differentiate between species' expressed gene complement barring genomic studies. Furthermore, by examining these potentially 'hardwired' differences between species' transcriptomic profiles we can elucidate the genetic basis for physiological divergences between dinoflagellate species.

Chapter 2

Saxitoxin Gene Structure and Representation in Non-toxic and Toxic Dinoflagellate Species

Abstract

Within microbial eukaryotes, dinoflagellates are among the most heavily impacted by horizontal gene transfer (HGT) events in their evolutionary history yet little is known about the fate of these sequences through the course of subsequent evolution and diversification. Here one example of a HGT event – the acquisition of saxitoxin biosynthesis genes - is examined across diverse saxitoxin producing and non-producing species. The cause of paralytic shellfish poisoning syndrome, the capability to produce saxitoxin was potentially acquired by dinoflagellates through HGT of saxitoxin biosynthesis genes from a cyanobacterium early in the radiation of peridinin-containing dinoflagellates. Currently, there are several genera of dinoflagellates that produce saxitoxin: *Alexandrium*, *Pyrodinium*, and *Gymnodinium*, but most species within and allied to these lineages do not. Here, deep transcriptome sequencing is used to examine the completeness of the saxitoxin biosynthesis pathway across these three genera as well as a related non-PST producing genus *Gambierdiscus* in order to document remnant sequences still present in this latter, non-producing genus. Through a BLAST-based approach, phylogenetic analyses were completed to determine dinoflagellate sequences homologous to genes within the cyanobacterial saxitoxin biosynthesis pathway. The first three genes to act in the saxitoxin biosynthesis pathway (sxtA, sxtB, and sxtG) are highlighted, as loss of any of these genes would potentially block toxin production before more cellular energy is lost in constructing intermediates. Finally, the completeness of the saxitoxin pathway is compared between toxic and non-toxic species in order to identify common pathways to loss of production and to provide further evidence for a single early HGT event. This approach reveals previously unknown diversity among sxtA among dinoflagellates that produce saxitoxin, many of which have orthologous sequences in *Gambierdiscus*. This result is further evidence that a single distant HGT event conferred saxitoxin biosynthesis genes from a cyanobacterium to the dinoflagellate ancestor of *Pyrodinium* and *Alexandrium*.

Introduction

Horizontal gene transfer (HGT) is increasingly recognized as a major driver of evolution and diversification among microbial eukaryotes (Keeling & Palmer, 2008; Soucy, Huang, & Gogarten, 2015). Though initially thought to be mainly important in prokaryotic evolution, high throughput genetic sequencing of protist species has shown that HGT events have occurred much more frequently than expected among eukaryotic species and often across domains of life, enabling the adaptation of eukaryotic organisms to new niches (Andersson, 2009; Raymond et al., 2012; Schönknecht, Weber, & Lercher, 2014). HGT events are also especially common in the evolution of dinoflagellates, far more common than in a cross section of other protists including stramenopiles, a haptophyte, a cryptophyte, prasinophytes, and other aveolates (Wisecaver et al., 2013).

Dinoflagellates share an evolutionary history that includes multiple endosymbiotic events and it has been proposed that HGT occurs mainly through endosymbiosis in eukaryotic organisms (Katz et al., 2015; Ku et al., 2015). This may have contributed to the importance of HGT events in their lineage. Furthermore, endosymbiosis may be an even greater driver of HGT in dinoflagellates since they have been shown to incorporate an unusually high proportion of plastid genes to their nucleus, more so than many other eukaryotes (Archibald, 2009; Dorrell & Howe, 2015). Dinoflagellates also 'recycle' genetic material to regulate their plastid photosynthetic

machinery from multiple sources which sometimes results in the fusion of domains from different organisms into chimeric sequences (Méheust et al., 2016). They also recycle genes for use in different pathways, resulting in at least one documented example of a hybrid metabolic pathway with components from different evolutionary origins (Bentlage, Rogers, Bachvaroff, & Delwiche, 2016). In addition, dinoflagellates frequently are hetero- or mixotrophic, allowing them to encounter more nucleic acids, thereby increasing opportunities for HGT events. Even dinoflagellates thought to be purely phototrophic have recently been discovered to engage in mixotrophy (Jacobson & Anderson, 1996; Jeong et al., 2010). Furthermore, their unusual practice of reverse transcription of mRNAs back into their DNA also facilitates incorporation of material into their nucleus (Slamovits & Keeling, 2008). All these attributes could potentially explain the particular influence of HGT on dinoflagellate evolution.

In dinoflagellates, HGT events have resulted in large-scale information transfer between species and across domains of life. This chapter examines one such example, a complex biosynthetic pathway responsible for the production of saxitoxins, which cause paralytic shellfish poisoning (PSP) when they contaminate shellfish and other seafood (Cembella, 1998). Saxitoxins are produced by both cyanobacteria and dinoflagellates, indicating at least one HGT event transferring this cluster of biosynthetic machinery between cyanobacteria and dinoflagellates. Interestingly, the production of saxitoxins within dinoflagellates is paraphyletic; saxitoxin producers include species within several unallied genera: *Alexandrium*, *Pyrodinium*, and *Gymnodinium.* Furthermore, within *Alexandrium* many species do not synthesize toxins. Therefore, if a single HGT event conferred the ability to produce saxitoxins from a cyanobacterial source, this ability must have been lost repeatedly across different dinoflagellate lineages (E.L. Lilly, Halanych, & Anderson, 2005; Orr et al., 2013a).

Research into the saxitoxin biosynthesis pathway has proceeded rapidly due to its economic and ecological significance as well as increased utility of molecular techniques, especially within cyanobacteria. Meticulous biochemical investigations established the chemical intermediates of the saxitoxin biosynthetic pathway (Shimizu et al., 1993, 1984). Gene function was then established though bioinformatic investigations of prokaryotes and the discovery of the saxitoxin gene cluster: 26 proteins encoded by genes sxtA-Z (Kellmann et al., 2008). Several homologous sequences from diverse saxitoxin-producing dinoflagellate species were then described, including several *A. tamarense* complex species, *Gymnodinium catenatum*, and *Pyrodinium bahamense* (Hackett et al., 2012; Stüken et al., 2011). Two sequences, a domain within the sxtA gene (sxtA4) and the sxtG gene, have thus far been found only in species that produce saxitoxin and not in closely related non-toxic dinoflagellate lineages (Murray et al., 2015; Orr et al., 2013b; Suikkanen, Kremp, Hautala, & Krock, 2013). Paralogous sequences to the sxtA gene have been reported as well; these sequences do not contain the sxtA4 domain and are found across a wider diversity of dinoflagellates (Murray et al., 2015; Stüken et al., 2011).

This study explores the spread and retention of genes involved in saxitoxin biosynthesis through examination of transcriptomes from a diverse set of dinoflagellates. Included in this set are PSP-causing species from the genera *Pyrodinium* and *Gymnodinium* as well as several species and isolates from the genus *Alexandrium*. Within *Alexandrium* the most widespread toxin producing species lie in the *A. tamarense* species complex subgroup (Emily L. Lilly et al., 2007). There are five recognized species within this complex, three of which produce saxitoxins (*A. fundyense*, *A. pacificum,* and *A. australiense*) and two that do not (*A. mediterraneum* and *A. tamarense*) (John et al., 2014), all of which

are represented in this analysis. Non-saxitoxin producing species from the genus *Gambierdiscus* are also included. The initial questions of this chapter are: do distantly related saxitoxin producers *Gymnodinium catenatum* and *Pyrodinium bahamense* contain all the same saxitoxin genes as *Alexandrium* and are there remnants of a primordial saxitoxin biosynthesis pathway in non-saxitoxin producing dinoflagellates? I show the pervasive necessity of sxtA4 and sxtG for saxitoxin biosynthesis in dinoflagellates and the retention of saxitoxin sequences in non-saxitoxin-synthesizing lineages. This leads directly to subsequent questions to be addressed: are there common patterns leading to loss of toxin synthetic capability and why are saxitoxin transcripts being conserved and expressed in nontoxic lineages? Through a BLAST-based approach, phylogenetic analyses were completed to determine dinoflagellate sequences homologous to genes within the cyanobacterial saxitoxin biosynthesis pathway. This chapter begins by examining the first three genes to act in the saxitoxin biosynthesis pathway, as loss of any of these genes would potentially block toxin production before more cellular energy is lost in constructing intermediates. Finally, the completeness of the saxitoxin pathway is examined between toxic and nontoxic species. The aim of these comparisons is to describe the differences between toxic and non-toxic species that are likely to underlie differences in toxin production and provide further evidence for an early HGT event.

Methods

Gene sequencing

Due to their large and chromosomally rich genomes, it remains impractical to sequence most dinoflagellate genomes. Deep transcriptome sequencing circumvents this problem by only characterizing transcribed genes. Here I present differences in transcriptomic gene composition as a proxy for genomic differences between dinoflagellate species. Sequence data were compiled from several sources including previously published studies, the Moore Microbial Eukaryote Sequencing Project, and new sequencing conducted during the present study (Table 2.1).

Illumina Sequencing and *de novo* Sequence Assembly

Cultures were grown and harvested during the exponential growth phase as described in Hackett et al., 2012 (Hackett et al., 2012). Bulk RNA was extracted through a Trizol extraction protocol and underwent poly-A selection. High throughput transcriptome sequencing (RNAseq) was performed using paired-end reads on the Illumina Genome Analyzer IIx Platform on four strains from three species within the *A. tamarense species complex*: *A. fundyense*, *A. tamarense*, and *A. pacificum* as described in Hackett et al., 2012. Strand-specific high throughput transcriptome sequence (ssRNAseq) was then performed with Illumina HiSeq on eleven additional strains from within the *A. tamarense* species complex: six toxic *A. fundyense* clones, one non-toxic *A. mediterraneum* clone, two non-toxic *A. tamarense* clones, one toxic *A. pacificum* clone and a toxic *A. australiense* clone. One strain of *A. ostenfeldii*, *A. affine*, *G. belizeanus, G. caribaeus, G. pacificus, G. silvae and G. sp. type 4* were also sequenced.

Four cDNA libraries were created from *A. fundyense* clones GTM253 and 38-3, two libraries from *A. tamarense* clone ATSP1-B, and four libraries from *A. pacificum* clone CCMP1598 (Hackett et al., 2012). Since there is currently no standard method for *de novo* transcriptome assembly these reads were assembled by multiple programs then clustered together to remove biases from any single

assembly method. The following four programs were used for *de novo* assembly of these sequencing reads: *CLC Genomics Workbench* (https://www.qiagenbioinformatics.com/), *Velvet/Oases* (Schulz, Zerbino, Vingron, & Birney, 2012), *trans-ABySS* (Robertson et al., 2010), and *Trinity* (Grabherr et al., 2011) . The output contigs from each program were then combined and clustered by *Cluster Database at High Identity with Tolerance* (CD-HIT) (Li & Godzik, 2006). In the CD-HIT analysis contigs were first combined then size-limited to 200 to 5000 base pairs. They were then clustered together with CD-HIT-EST in a stepwise fashion to 95% homology over 80% of their lengths. Based on this sequencing analysis, we chose to proceed with de novo assembly using *Trinity* for the second group of ssRNAseq libraries. No contig clustering was then performed for the output of these latter assemblies. Each contig library was also run through the *Benchmarking Universal Single-Copy Orthologs* (BUSCO) software package to asses the transcriptome's completeness (Simão et al., 2015).

Species	Strain	Source	# Contigs	# Trimmed Reads	N50
Alexandrium affine	PA5V	This Study	112174	115092316	382
Alexandrium andersoni	CCMP2222	MMETSP	42240		735
Alexandrium australiense	ATBB01	This Study	76161	50202246	357
Alexandrium fundyense	CCMP1719	MMETSP	11897		253
Alexandrium fundyense	UW462	This Study	26741	165505834	297
Alexandrium fundyense	ARBR2G	This Study	41757	177518388	332
Alexandrium fundyense	SPE10	This Study	72688	48028494	360
Alexandrium fundyense	PW06	This Study	87477	50896092	367
Alexandrium fundyense	At383	Hackett et al., 2012	100608	53478181	1007
Alexandrium fundyense	AtGTM253	Hackett et al., 2012	136077	68594660	1207
Alexandrium margalefi	AMGDE01CS-322	MMETSP	54023		1055
Alexandrium mediterraneum	SZB	This Study	32302	191049824	304
Alexandrium minutum	CCMP113	MMETSP	13537		810
Alexandrium monilatum	CCMP3105	MMETSP	99594		1351
Alexandrium ostenfeldii	LK_ost	This Study	11803	191902822	307
Alexandrium pacificum	ACPP01	This Study	48395	199087382	325
Alexandrium pacificum	OF101	MMETSP	72232		1183
Alexandrium pacificum	At1598	Hackett et al., 2012	128630	136820044	1352
Alexandrium tamarense	ATSPF7	This Study	26866	162080468	293
Alexandrium tamarense	ATSW01	This Study	27791	90974774	312
Alexandrium tamarense	UW452	This Study	107653	101561860	378
Alexandrium tamarense	CCMP 1771	MMETSP	135584		1364
Alexandrium tamarense	AtSP1B	Hackett et al., 2012	147548	67370976	1226
Gambierdiscus australes	CAWD 149	MMETSP	53551		1201
Gambierdiscus belizeanus	FCMay10_8	This Study	146019	40518428	447
Gambierdiscus caribaeus	BPAug08	This Study	107034	21170278	406
Gambierdiscus pacificus	D50511-03	This Study	134248	25201812	458
Gambierdiscus silvae	FCMay10_9	This Study	19750	89124466	317
Gambierdiscus species type 4	ID00-07	This Study	138533	85201912	440
Gymnodinium catenatum	GC744	MMETSP	88811		1621
Pyrodinium bahamense	pbaha01	MMETSP	105175		1458

Table 2.1 Transcriptome Libraries
Species and Strain name, origin of data, and number of contigs for each transcriptome library examined. Number of trimmed reads given for transcriptomes assembled in this study as well as N50 statistic for all transcriptome libraries.

Basic Local Alignment Search Tool (BLAST) Analysis

Custom BLAST databases were created out of each output of RNAseq clustered contigs (*A. fundyense*, *A. tamarense*, and *A. pacificum)* as well as ssRNAseq Trinity assemblies (*A. fundyense*, *A. mediterraneum*, *A. tamarense*, *A. pacificum*, *A. australiense*, *A. ostenfeldii*, *A. affine*, *G. belizeanus, G. caribaeus, G. pacificus, G. silvae and G. sp. type 4)* and on assemblies accessed from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) funded by the Moore Foundation (*A. fundyense, A. monilatum, A. tamarense, A. minutum, A. margalefii, A. pacificum, A. andersoni, G. australes, Gymnodinium catenatum,* and *Pyrodinium bahamense*). Using both cyanobacterial saxitoxin synthesis

gene sequences and dinoflagellate sequences that have been defined by previous research to be homologous to cyanobacterial saxitoxin genes, blastn and tblastn searches were conducted against the custom BLAST databases. Query sequences included previously discovered cyanobacterial saxitoxin sequences as well as two sxtA dinoflagellate isoforms (Stüken et al., 2011); six dinoflagellate sxtA sequences (Hackett et al., 2012); two dinoflagellate sequences of sxtB (Hackett et al., 2012); and four dinoflagellate sequences of sxtG (Hackett et al., 2012; Orr et al., 2013b).

SxtA, B, G Analysis

For detailed analysis of the first three genes to act in the saxitoxin biosynthesis pathway, sxtA, B, and G, contigs with query hits with an e-value less than 10E-5 (0.00001) were selected for analysis. The contigs' reciprocal BLAST hit was then found in blastx searches to NCBI's nr database. Only those with significant hits back to a domain found within the cyanobacterial saxitoxin genes were kept. Contigs were translated into their correct frame according to their tblastn hit against the protein query then aligned using *Geneious* software (Kearse et al., 2012) and manually checked for errors. Maximum Likelihood analysis was completed using *RAxML* optimizing for ML model and number of bootstraps (Stamatakis, 2014). Within *Geneious* the program *InterProScan* was used to predict each contig's protein domains (Li & Godzik, 2006; Zdobnov & Apweiler, 2001).

Complete Saxitoxin Gene Pathway Analysis

To analyze the entire saxitoxin biosynthesis pathway cyanobacterial saxitoxin genes A-Z were queried against the transcriptome data set. All hits with an evalue less than 10E-5 were selected for analysis. The contigs' reciprocal BLAST hit was then found in blastx searches to NCBI's nr database. Contigs were translated into their correct frame according to their tblastn hit against the protein query then aligned using *MSAprobs* (Yongchao et al., 2010). Maximum Likelihood analysis was completed using *RAxML* optimizing for ML model and number of bootstraps (Stamatakis, 2014).

Supplemental Table 2.1: Queries used in BLAST Analysis

Open Reading Frame Prediction, Translation, BUSCO and KEGG Analysis

To translate contigs into their most likely protein sequence open reading frame sequences of the contig libraries were predicted and translated using *OrfPredictor* (Min, Butler, Storms, & Tsang, 2005). As input a blastx results file of the contig library verses NCBI's nr database was provided in addition to the library's sequence fasta file to improve prediction accuracy. The resulting output of predicted amino acid sequences were then fed into the *Kyoto Encyclopedia of Genes and Genomes'* (KEGG) online Ghost Koala interface (Kanehisa, Sato, Kawashima, Furumichi, & Tanabe, 2016). This analysis was completed for fourteen transcriptomes in total, at least one from each species of the *A. tamarense* species complex.

Supplemental Figure 2.1: KEGG Analysis of Transcriptome Libraries

Results

Species	Strain ID	Source	Complete	Duplicate	Fragmented	Missing
A. tamarense	CCMP1771	MMETSP	67	34	9.5	22
G. catenatum	GC744	MMETSP	65	27	5.1	28
A. monilatum	CCMP3105	MMETSP	63	14	7.4	28
A. pacificum	At1598	Hackett et al., 2012	62	19	9	28
A. tamarense	AtSP1B	Hackett et al., 2012	61	17	10	28
A. fundyense	AtGTM253	Hackett et al., 2012	61	17	8.8	29
P. bahamense	pbaha01	MMETSP	58	22	9.5	31
A. pacificum	OF101	MMETSP	51	16	14	34
A. fundyense	At383	Hackett et al., 2012	45	12	18	35
A. margalefi	AMGDE01CS-322	MMETSP	45	12	11	43
G. australes	CAWD 149	MMETSP	44	8.8	12	43
G. pacificus	D50511-03	This Study	29	6.2	20	50
G. belizeanus	FCMay10_8	This Study	27	5.8	23	48
G. sp. type 4	ID00-07	This Study	25	5.5	24	50
G. caribaeus	BPAug08	This Study	20	4.1	21	57
A. andersenii	CCMP2222	MMETSP	19	3.7	15	65
A. minutum	CCMP113	MMETSP	14	3.7	4.4	80
A. tamarense	UW452	This Study	12	2.7	25	62
A. australiense	ATBB01	This Study	11	3	21	66
A. fundyense	SPE10	This Study	11	1.6	20	67
A. affine	PA5V	This Study	11	1.3	19	68
A. fundyense	PW06	This Study	10	3.7	24	64
A. tamarense	ATSW01	This Study	6.5	1.1	12	80
A. fundyense	ARBR2G	This Study	4.8	1.1	14	80
G. silvae	FCMay10_9	This Study	3.9	0.4	10	86
A. fundyense	UW462	This Study	3.7	0.2	10	85
A. mediterraneum	SZB	This Study	3.4	0.4	8.8	87
A. pacificum	ACPP01	This Study	2.5	0.2	10	86
A. tamarense	ATSPF7	This Study	2.5	0	9	88
A. ostenfeldii	LK_ost	This Study	2.3	0	7.9	89
A. fundyense	CCMP1719	MMETSP	2.3	0.6	7.4	90

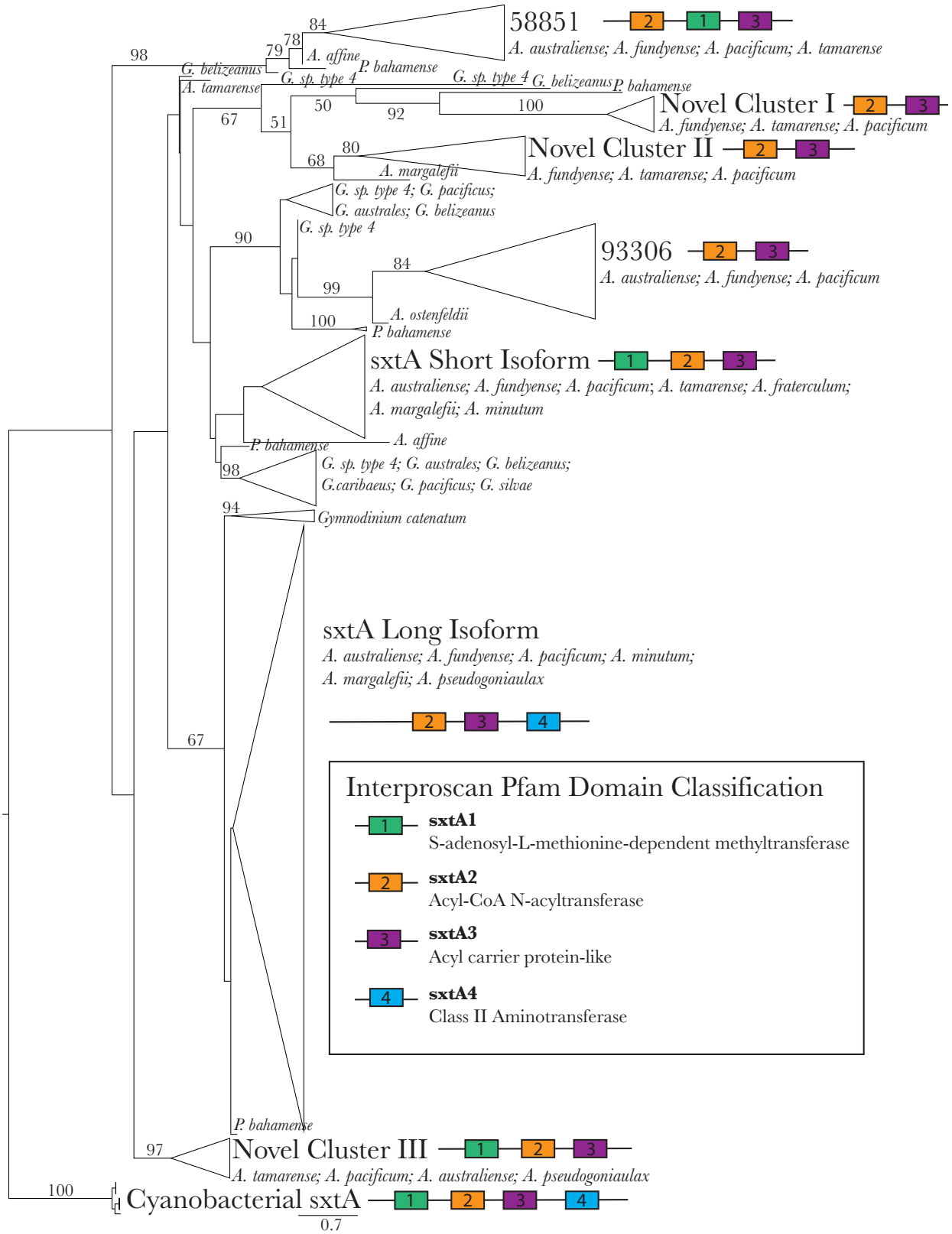
Table 2.2 BUSCO Results
Transcriptome libraries were queried with the program *Benchmarking Universal Single-Copy Orthologs* (BUSCO) that determines a transcriptome's completeness though the presence of orthologous groups common to all eukaryotes. It then gives the percent complete, fragmented, duplicated, and missing of these groups.

Transcriptome Quality

Transcriptome quality ranged from relatively high quality libraries with *Benchmarking Universal Single-Copy Orthologs* (BUSCO) percent complete groups scores ranging from 67% to 58% (*A. fundyense*, *A. monilatum*, *A. pacificum*, *A. tamarense*, *G. catenatum*, and *P. bahamense*; Table 2.2) to libraries with especially low scores (ten libraries with ≤10% complete BUSCO groups; Table 2.2). Relatively high BUSCO-scoring transcriptome libraries also have higher N50 scores (Table 2.1), supporting their higher quality. In subsequent analysis, presence and absence of transcripts

in species with relatively high scoring libraries (listed above) will be considered more significant than species with lower-scoring transciptome libraries. Absence of transcipts in especially low scoring libraries (≤10% complete BUSCO groups) will be considered likely due to poor transcriptome quality. All transcriptome libraries are not a complete representation of coding sequences present in these species, however, higher quality libraries are representative of transcribed sequences in their source cultures.

Figure 2.1 Phylogenetic tree of sxtA sequences
Phylogenetic analysis of sxtA BLAST hits. *RAxML* was run on a protein alignment of 313 sequences. Support values represent ML bootstrap values (only those values above 50 are represented). *Interproscan* predicted domains are pictured in numbered boxes. BLAST result contigs were included as well as two sxtA isoforms from (Stüken et al., 2011), four sxtA homologs from Hackett et al., 2012, and sxtA1 and sxtA4 domain sequences from Murray et al., 2015.


58851
A. australiense; A. fundyense; A. pacificum; A. tamarense
Novel Cluster I
A. fundyense; A. tamarense; A. pacificum
Novel Cluster II
A. fundyense; A. tamarense; A. pacificum
93306
A. australiense; A. fundyense; A. pacificum
sxtA Short Isoform
A. australiense; A. fundyense; A. pacificum; A. tamarense; A. fraterculum;
A. margalefii; A. minutum
sxtA Long Isoform
A. australiense; A. fundyense; A. pacificum; A. minutum;
A. margalefii; A. pseudogoniaulax
Novel Cluster III
A. tamarense; A. pacificum; A. australiense; A. pseudogoniaulax
Cyanobacterial sxtA
0.7
A. affine
P. bahamense
G. belizeanus
A. tamarense
G. sp. type 4
G. sp. type 4 G. belizeanus
P. bahamense
A. margalefii
G. sp. type 4; G. pacificus;
G. australes; G. belizeanus
G. sp. type 4
A. ostenfeldii
P. bahamense
P. bahamense
A. affine
G. sp. type 4; G. australes; G. belizeanus;
G.caribaeus; G. pacificus; G. silvae
Gymnodinium catenatum
P. bahamense
Interproscan Pfam Domain Classification
sxtA1
S-adenosyl-L-methionine-dependent methyltransferase
sxtA2
Acyl-CoA N-acyltransferase
sxtA3
Acyl carrier protein-like
sxtA4
Class II Aminotransferase

SxtA Analysis

Searches for sxtA in the compiled transcriptome assemblies recovered seven distinct clusters of contigs that are similar to cyanobacterial sxtA (Figure 2.1). Four clusters correspond to previously described sequences from studies by Hackett et al (2012) and Stüken et al (2011) while three clusters potentially represent new variants of sxtA (Hackett et al., 2012; Stüken et al., 2011). InterProScan identified protein domains for a SAM-dependent methyltransferase, a GCN5-related N-acyltransferase, an acyl carrier protein and a class II aminotransferase corresponding to sxtA1-4 respectively. Two clusters of sequences associated with the short isoform identified by Stüken et al (2011), one cluster associated with the sxtA homolog identified by Hackett et al (2012), and one novel cluster contain predicted domains for sxtA1-3. Those sequences clustering with the long isoform identified by Stüken et al (2011) mostly contained domains for sxtA2-4 (except genomic PCR transcripts from the Murray et al., 2015 study which were predicted to contain sxtA1). Three of the remaining clusters of sequences, two novel and one associated with a sxtA homolog identified by Hackett et al (2012), contain domains for sxtA2 and 3.

SxtG Analysis

Sequences orthologous to the dinoflagellate sxtG clade were found only in toxic strains. However, several other clades of sequences similar to sxtG were found across the dinoflagellates examined, all containing domains predicted to encode glycine amidinotransferases (Figure 2.2). This phylogeny is similar to that described in Murray et al., 2015 across a wide array of dinoflagellate species (Murray et al., 2015). These paralogs may function in other cellular pathways since they are conserved among many *Alexandrium* strains in this analysis.

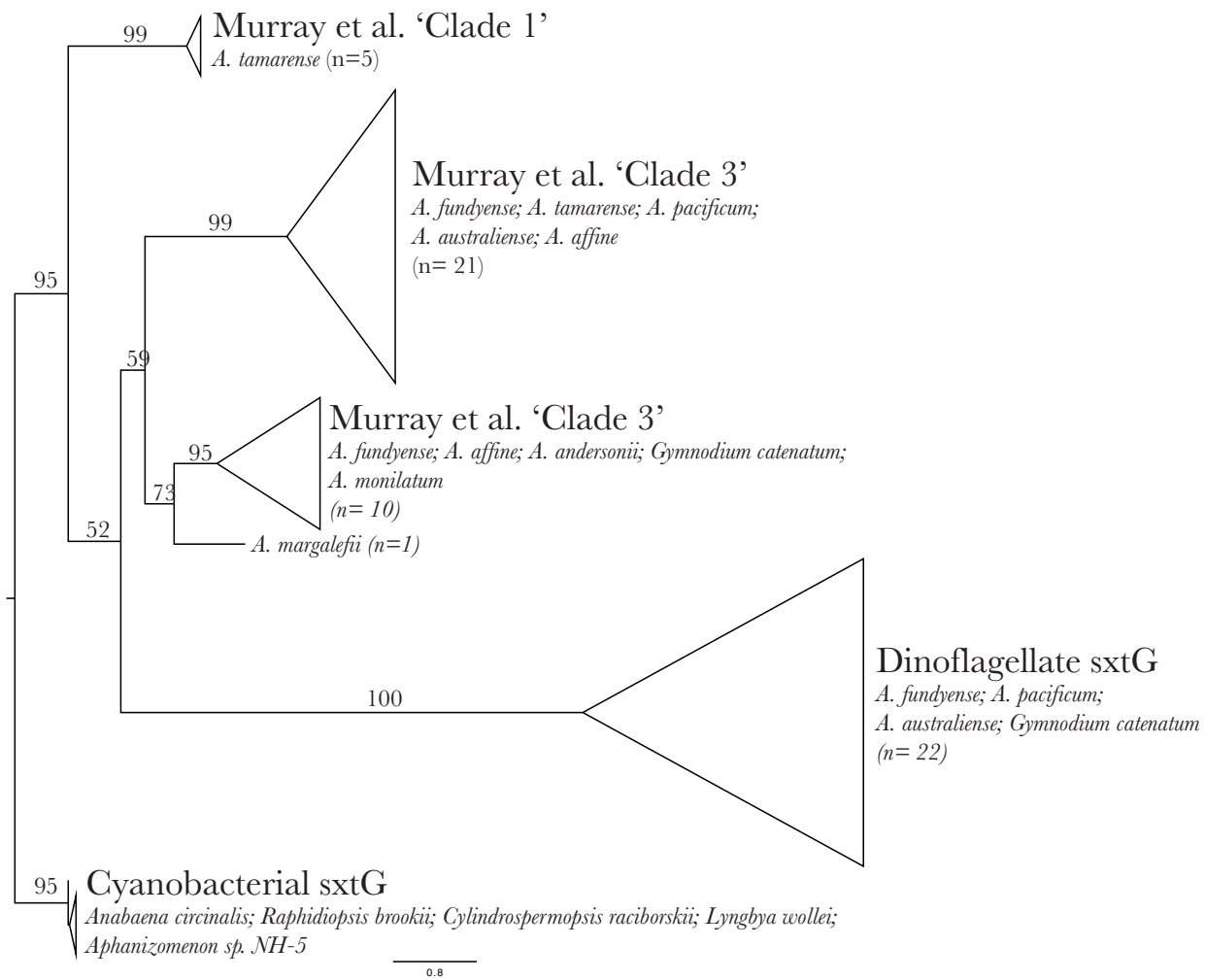


Figure 2.2 Phylogenetic Tree of sxtG BLAST Hits
Phylogenetic analysis of sxtG BLAST hits. *RAxML* was run on a protein alignment of 65 sequences. Support values represent ML bootstrap values (only those values above 50 are represented).

SxtB Analysis

Both toxic and non-toxic strains contained similar sequences to the dinoflagellate sxtB homologs as identified by Hackett et al. (Figure 2.3) (Hackett et al., 2012). *InterProScan* predicted cytidine deaminase domains in both clades. No contigs similar to sxtB were found in *Gymnodinium catenatum* or *Pyrodinium bahamense* transcriptomes.

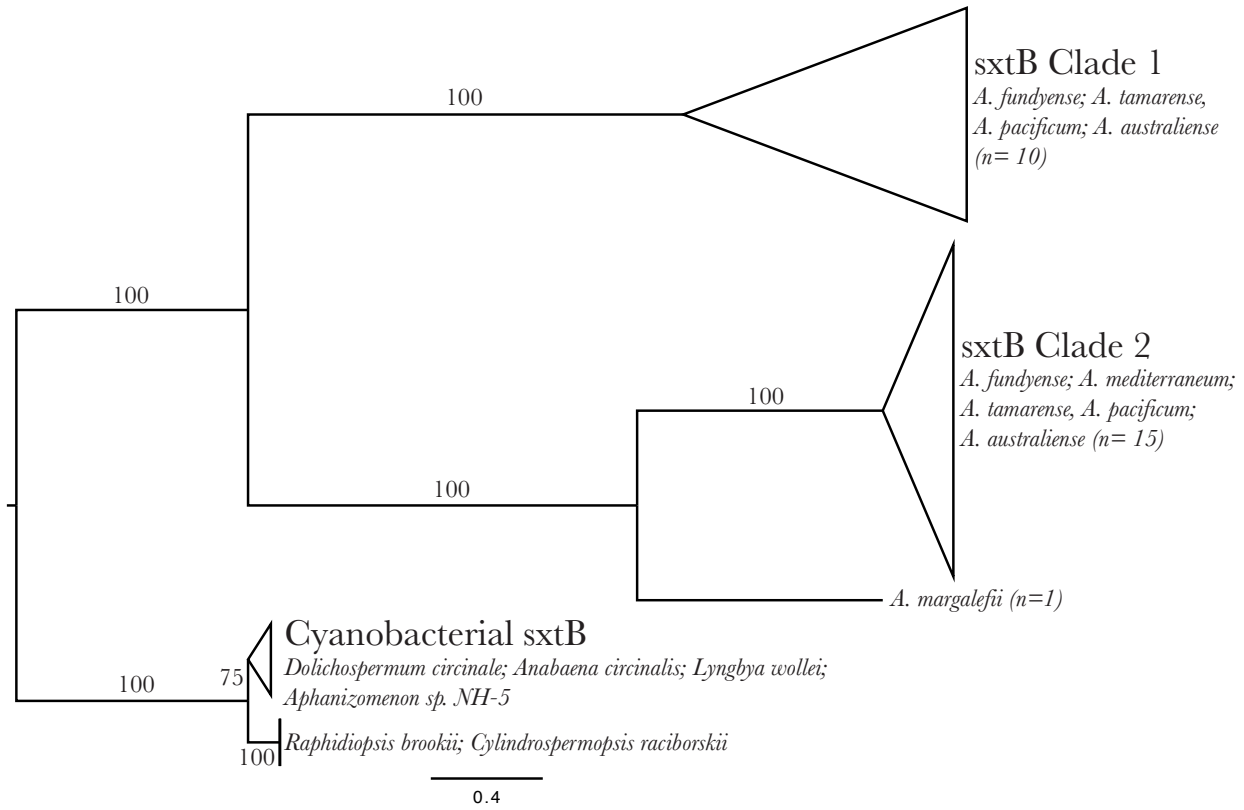


Figure 2.3 Phylogenetic Tree of sxtB BLAST hits
Phylogenetic analysis of sxtB BLAST hits. *RAxML* was run on a protein alignment of 33 sequences. Support values represent ML bootstrap values (only those values above 50 are represented).

Further Saxitoxin Gene Analysis

Cyanobacterial saxitoxin genes A-Z were queried against the transcriptome data set. Hits were aligned with *MSAprobs* and phylogenetic analysis was conducted with *RAxML*. All hits with an evalue less than 10E-5 were kept. This analysis did not include manual curation of alignments based on reciprocal BLAST nr hit, as with sxtA, B, and G genes. All hits were selected based purely on evalue score. All saxitoxin genes except sxtC, E, J, K, and R were found in at least one toxic *Alexandrium tamarense* species complex strain (Table 2.3). Eleven out of fourteen (11/14) of the saxitoxin 'core' genes set (conserved in five species of saxitoxin producing cyanobacteria) were found and of this set all eight of the genes predicted to be directly involved in saxitoxin synthesis in cyanobacteria had at least one predicted *Alexandrium tamarense* sp. complex homolog. All three genes involved in modifying saxitoxin into its congeners were also found (sxtL, N, X). In both this study and the study by Hackett et al., 2012, dinoflagellate genes relating to sxtC, E, J, K, and R remain undescribed. However, in the study by Stüken et al., 2011, the authors report finding sxtR in an *A. fundyense* strain.

			Saxitoxin Biosynthesis Pathway Genes																										
Species	Strain	STX+	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A. australiense	ATBB01	Yes	13	3				2	2	2				1	2	1	1	1				1	1		1			1
A. fundyense	At383	Yes	15	4				3	5	8	4				4	1	3	1				8	43	1	1	2		9
A. fundyense	GTM253	Yes	23	3				3	11	13	2				3	1	3	2			1	12	55	1	2	2		10
A. fundyense	SPE10	Yes	13	4					4	3	1							1	1			1	1	2				
A. fundyense	ARBR2G	Yes	6						4					1				1					1					
A. fundyense	CCMP1719	Yes														2							3					
A. fundyense	UW462	Yes	6								1			1														
A. fundyense	PW06	Yes	13						3	5	1			1				1	1			1	6	1				
A. pacificum	At1598	Yes	30	3		1		3	2	16	4				3	2	5	4				15	83		2	1	1	21
A. pacificum	ACPP01	Yes		1					1																			
A. pacificum	OF101	Yes								1					1	1						2				1		
A. mediterranei	SZB	No		1										1									3					
A. tamarense	CCMP1771	No	4	3					2		1				3	2	10	1				15				1		12
A. tamarense	CCMP1771	No	15	2				2	2	15				1	1	2		1					52			1		17
A. tamarense	AtSP1B	No	23	2				4	2	15				1	4	2	4	2				14	59		2	1		10
A. tamarense	UW452	No	10	2				2	5	3				1		1	1					3	8		2	1		1
A. tamarense	ATSW01	No	3																				2		1	1		
A. tamarense	ATSPF7	No	6						3																			
A. ostenfeldii	LK_ost	Yes	2														1											1
A. minutum	CCMP113	Yes																					9					
A. affine	PA5V	No	3	1					3	4								1				1	11			1		
A. andersoni	CCMP2222	No							1							1									1			1
A. margalefi	AMGDE01CS-322	No	1	1					1	5					1							1	55	1	3	1		2
A. monilatum	CCMP3105	No						2	2	22						2	2					4	56		2	1		
G. australes	CAWD 149	No	7					3		3					6	1						3	24		1			
G. belizeanus	FCMay10_8	No	11					2		1					2		1				2	1	19		5			3
G. caribaeus	BPAug08	No	4													1	3	1					15		5			1
G. pacificus	D50511-03	No	6					1		2					2	1	1					2	1		1			
G. silvae	FCMay10_9	No	1					1									1						1		1			
G. sp. type 4	ID00-07	No	10					1		1					4		1					2	15					
G. catenatum	GC744	Yes	4					3	2	10	2				3	4	7					7				3		9
P. bahamense	pbaha01	Yes	12					1			1				5	2	2					1	2			4		
		Total	241	30	0	1	0	33	55	129	17	0	0	8	44	27	46	17	2	0	3	94	525	6	30	21	1	98

Table 2.3 Number of Contigs Similar to each Saxitoxin Biosynthesis Pathway Gene per Library

Shows the number of contigs per library that passed similarity thresholds to saxitoxin biosynthesis genes. Libraries are organized by larger grouping (membership within *A. tamarense* species complex, genera, and toxin production.

Discussion

Conservation of saxitoxin biosynthesis pathway genes across saxitoxin-producing genera

The genera *Gymnodinium*, *Pyrodinium* and *Alexandrium* all contain species that produce saxitoxins, but while *Pyrodinium* and *Alexandrium* have a relatively recent common ancestor, *Gymnodinium* diverged much further in the past (Murray et al., 2015). One of the aims of this chapter is to determine whether saxitoxin producers within these genera contain a similar range of saxitoxin biosynthetic transcripts or if there is a greater level of divergence between *Alexandrium* and *Gymnodinium*, than *Alexandrium* and *Pyrodinium*. If a lone HGT event conferred the ability to produce saxitoxins from a cyanobacterium on a single common ancestor of all three of these

genera, transcripts within *Gymnodinium* would be more divergent. This study found eight saxitoxin genes that were expressed in at least one species within *Alexandrium*, *Gymnodinium catenatum*, and *Pyrodinium bahamense* (AFIMNOTX). Nine genes have hits from *Alexandrium* and *Pyrodinium bahamense* (AFIMNOTUX) and eleven genes between *Gymnodinium catenatum* and *Alexandrium* (AFGHIMNOTXZ). Therefore, this study found more genes shared between *G. catenatum* and *Alexandrium* rather than *P. bahamense* and *Alexandrium.* This could be indicative of a process that caused saxitoxin genes in *P. bahamense* to either diverge beyond the limits of detection of this method or for *P. bahamense* to substitute genes from another origin for functions required in saxitoxin synthesis. However, within the cluster of sxtA transcripts that contained sequences from *Alexandrium*, *G. catenatum*, and *P. bahamense* ('sxtA Long Isoform') transcripts maintained the same phylogenetic relationship as trees constructed from ribosomal sequences: *G. catenatum* sequences were basal to *P. bahamense* and *Alexandrium* sequences (Figure 2.1). These discrepancies could also be due to the quality of the *G. catenatum* verses *P. bahamense* transcriptomes: the transcriptome of *G. catenatum* has a moderately higher BUSCO score than that of *P. bahamense*. However, both transcriptome libraries are of relatively high quality and so likely represent their cultures' expressed transcripts.

The saxitoxin genes that had no hits in this study (Table 2.3, sxtC, E, J, K, and R) had functions predicted to be 'regulatory', 'unknown', 'regulatory', 'regulatory', and 'unknown' respectively (Kellmann et al., 2008). When dinoflagellates incorporated cyanobacterial genes it is possible they put them under dinoflagellate-specific regulation and had less need for these cyanobacterial regulatory elements.

Representation of saxitoxin biosynthesis pathway across non-saxitoxin producing dinoflagellates

Consistent with previous research, I did not find transcripts for sxtA4 within any non-toxic species. Similarly, a clade of sxtG containing only toxic members remains likely to contain the true homologous sequences for the cyanobacterial sxtG in dinoflagellates.

The genus *Gambierdiscus* shares a common ancestor with saxitoxin-producing genera *Alexandrium* and *Pyrodinium*, but does not contain any species that produce saxitoxins. However, the same HGT event that transferred saxitoxin genes to the common ancestor of *Alexandrium* and *Pyrodinium* would have potentially left remnant genes behind in *Gambierdiscus*. In total, twelve out of twenty-six (12/26) saxitoxin genes had hits in *Gambierdiscus*, seven of which are in common with *Alexandrium* and *Pyrodinium bahamense* (AFMNOTU). *Gambierdiscus* itself is a potent toxin producer, producing maitotoxin, an extremely potent natural non-biopolymer, and gambiertoxins, the precursors to ciguatoxins that are responsible for ciguatera fish poisoning, a significant human health threat (see chapters 3,4). These cyanobacterial saxitoxin genes may have been repurposed to fill another cellular function within *Gambierdiscus* leading to their retention and expression. Genes not present in *Gambierdiscus* (BCDEGIJKLQRVXY) ranged over a variety of functions. Most surprising was the absence of sxtG since paralogs similar to this gene are present across many dinoflagellate lineages (Murray et al., 2015).

Within the *Alexandrium tamarense species complex,* this study also found saxitoxin transcripts that were expressed in at least one toxic species but not in any non-toxic species. There were six genes from this analysis that had hits from saxitoxin-producing *A. tamarense sp. complex* strains but no corresponding hits from non-toxic *A. tamarense sp. complex* strains (D,G,Q,S,V,Y). This is in

addition to sxtA4. These genes might be specific to the formation of saxitoxin as five out of six had no corresponding hits from *Gambierdiscus*. Furthermore, *A. affine*, a closely related non-toxic species, also does not have hits to any of these genes. The exception is sxtS that has two hits from *Gambierdiscus belizeanus*. In total, the non-toxic *Alexandrium* species *A. mediterraneum* and *A. tamarense* collectively contained 16/26 saxitoxin transcripts, though only three (BLU) were found in the single *A. mediterraneum* library, likely due to its poor transcriptome quality (Table 2.2).

Are there common patterns leading to loss of toxin synthetic capability?

Consistent with earlier reports, the sxtA long isoform identified in Stüken et al 2011 (the only orthologous group which contains the domain sxtA4) was not present across all non-toxic strains and was present across all toxic *Alexandrium tamarense* species complex strains (including an A. *australiense* strain, reinforcing the results from Murray et al., 2015 (Murray et al., 2015). Furthermore, a clade of sxtG previously reported to contain the cyanobacterial sxtG homologs contained only transcripts from toxic species. As the most recent species to potentially lose the capability to produce saxitoxin, the non-toxic species within the *Alexandrium tamarense sp. complex* were also missing transcripts for five additional genes (D,Q,S,V,Y) that were present in at least one toxic *Alexandrium tamarense sp. complex* species. However, these genes were also not present in *P. bahamense* and *G. catenatum*. These genes were found in only one out of three toxic species, either *A. fundyense* or *A. pacificum* depending on the gene, and were missing from *A. australiense*; therefore they could be expressed at low levels and were commonly missed or genes from other sources may have replaced some of sxtD,Q,S,V,Y functions. This illustrates one of the limitations with working with transcriptome data as it is difficult to differentiate between absence and low abundance of a transcript.

Why are saxitoxin transcripts being conserved and expressed in non-toxic lineages?

Dinoflagellates are able to recycle genes for use in alternate pathways. This study found evidence for this process occurring in the saxitoxin biosynthesis pathway. In the analysis of the sxtA gene, seven groups of sequences containing domains with a high degree of homology to those found in cyanobacterial sxtA were observed. Four of these groups have been previously defined but three are new to this analysis (Hackett et al., 2012; Murray et al., 2015; Stüken et al., 2011). They contain a similar protein structure to previously identified sxtA variants. In cyanobacteria, all four domains of sxtA are transcribed as a single gene. In dinoflagellates the sxtA gene's multiple domains (and therefore functions) have been split up within some of the groups of contigs. Variants of the domains sxtA1-3 were found in all five species of the *A. tamarense* species complex, however, sequences similar to the domain sxtA4 were only found in toxic strains. Assembled contigs from non-toxic strains contained no sequences that clustered with the long isoform from the Stüken et al (2011) study.

In Murray et al. (2015), the authors reported three clades of sequences similar to sxtA1. Clade 2, comprised of toxic strains, likely corresponds to the clade grouping with the sxtAlong isoform in this study. Clade 3, which was present across all dinoflagellates, includes members from the clade grouping with sxtA short and 93306. Clade 1, which mainly corresponded to sequences from non-toxic *Alexandrium* lineages, is similar to the novel cluster 3 from this study since several *A. pseudogoniaulax* sequences from that study clustered in the grouping. However, this study also

found sequences from toxic *A. pacificum* and *A. australiense* within novel cluster 3 which were absent in the Murray et al., 2015 study's Clade 1.

This study is presenting evidence of an increased diversity of dinoflagellate contigs similar to sxtA described from cyanobacteria, many of them expressed in non-saxitoxin producing lineages either within *Alexandrium*, or in *Gambierdiscus*. Clearly duplication and diversification of sxtA transcripts has led to a gene family which likely has functions outside of toxin synthesis, preserving the gene's expression in species which no longer produce saxitoxins. Further characterization of these alternate functions could shed light on why saxitoxin biosynthesis genes were originally retained within dinoflagellates. Although for a long time the ecological advantage of saxitoxin production in dinoflagellates was debated, more recently evidence has accumulated showing saxitoxin production may confer a significant advantage in reducing copepod grazing pressure (Lasley-rasher, Nagel, Angra, & Yen, 2016; Selander et al., 2015). An important question therefore is why saxitoxin production is occurring in some dinoflagellates and not others, specifically what is different about the ecological pressures between habitats of toxic and non-toxic species within the *A. tamarense* sp. complex? Answering these questions could also elucidate why the ability to produce saxitoxin has been selectively lost across many dinoflagellate lineages.

Evolutionary Implications of *Gambierdiscus* sequences

It has been hypothesized that the saxitoxin biosynthesis pathway was transferred from a cyanobacterium to a common ancestor of *Pyrodinium bahamense* and *Alexandrium* in a massive horizontal gene transfer (HGT) event (Murray et al., 2015; Orr et al., 2013a). If this occurred, we would expect to find remnants of these genes even in non-saxitoxin producing lineages that share this common ancestor. In this analysis, I've found sequences in *Gambierdiscus* similar to these cyanobacterial saxitoxin genes even though *Gambierdiscus* does not produce saxitoxin. As *Gambierdiscus* has been shown to share a common ancestor with *Alexandrium* and *Pyrodinium* (Murray et al., 2015) this presents further evidence for an ancestral HGT event.

In total there are similar sequences from *Gambierdiscus* for five of the clusters of sxtA contigs. Interestingly, *Gambierdiscus* wasn't found to transcribe any genes that clustered with the sxtA Long isoform that contains the sxtA4 domain. Instead, in addition to toxic *A. tamarense* complex species we found sxtA Long isoform transcripts from *Gymnodinium catenatum* and *Pyrodinium bahamense*, both saxitoxin producers. Since *G. catenatum* is more distantly related to *Alexandrium*, *Gambierdiscus*, and *Pyrodinium* and yet contains saxitoxin gene sequences, it's been hypothesized that the saxitoxin genes found in *G. catenatum* may be the result of a secondary dinoflagellate to dinoflagellate HGT event (Orr et al., 2013a) or that the initial HGT event occurred even more distantly to the common ancestor of *G. catenatum* and *Alexandrium*, and was subsequently lost over much of the subsequent dinoflagellate lineages (Murray et al., 2015). In order to differentiate between these two hypotheses, further analyses need to be done comparing saxitoxin gene complements across dinoflagellates more distant from *Pyrodinium*, *Alexandrium*, and *Gambierdiscus*. By expanding the characterization of saxitoxin sequences and diversity across a wider range of extant dinoflagellates, specifically those groups basal to *G. catenatum* and additionally those groups linking *G. catenatum* to *Alexandrium*, *Gambierdiscus*, and *Pyrodinium*, timing of these potential HGT events could be resolved.

Conclusion

By examining relationships among the *A. tamarense* species complex and their complement of saxitoxin genes, we can establish differences between the toxic and non-toxic species, illustrating genetic changes that may be important for toxin biosynthesis and further elaborating on the evolutionary history of the saxitoxin gene pathway in dinoflagellates. This chapter reports an increased diversity in sxtA among dinoflagellates that produce saxitoxin. Furthermore, many of these clusters are closely paired with *Gambierdiscus* sequences. This could be further evidence of a distant HGT event where components of the saxitoxin biosynthesis pathway were transferred from a cyanobacterium to a dinoflagellate ancestor of *Pyrodinium* and *Alexandrium*. This study also demonstrates the ability of dinoflagellates to recycle genes from multiple sources to accomplish different tasks, as conservation and expression of some saxitoxin genes in multiple *Gambierdiscus* species suggests their use in other pathways. Characterization of the function of these sequences in *Gambierdiscus* and other non-saxitoxin producing species could not only indicate their current utility but also why these cyanobacterial sequences may have been originally retained in a dinoflagellate ancestor. Furthermore, this analysis reinforces the findings that sxtA4 and one clade of sxtG are only found in toxic species and are therefore predicted to be necessary for toxin production. Overall, this chapter illustrates the genetic diversity that occurs through gene duplication and diversification of a biosynthesis pathway in dinoflagellates.

Funding Sources:
USFDA (F223201000060C)
NOAA NOS (CiguaHAB; NA11NOS4780060 and NA11NOS4780028);
COHH - Woods Hole Center for Oceans and Human Health – NSF Grant OCE-1314642 and NIEHS 1P01-ES021923-01
WHOI Ocean Ventures Fund; WHOI Coastal Institute Research Proposal Award

Chapter 3

Development of Fluorescence in situ Hybridization (FISH) Probes to Detect and Enumerate *Gambierdiscus* Species

Abstract

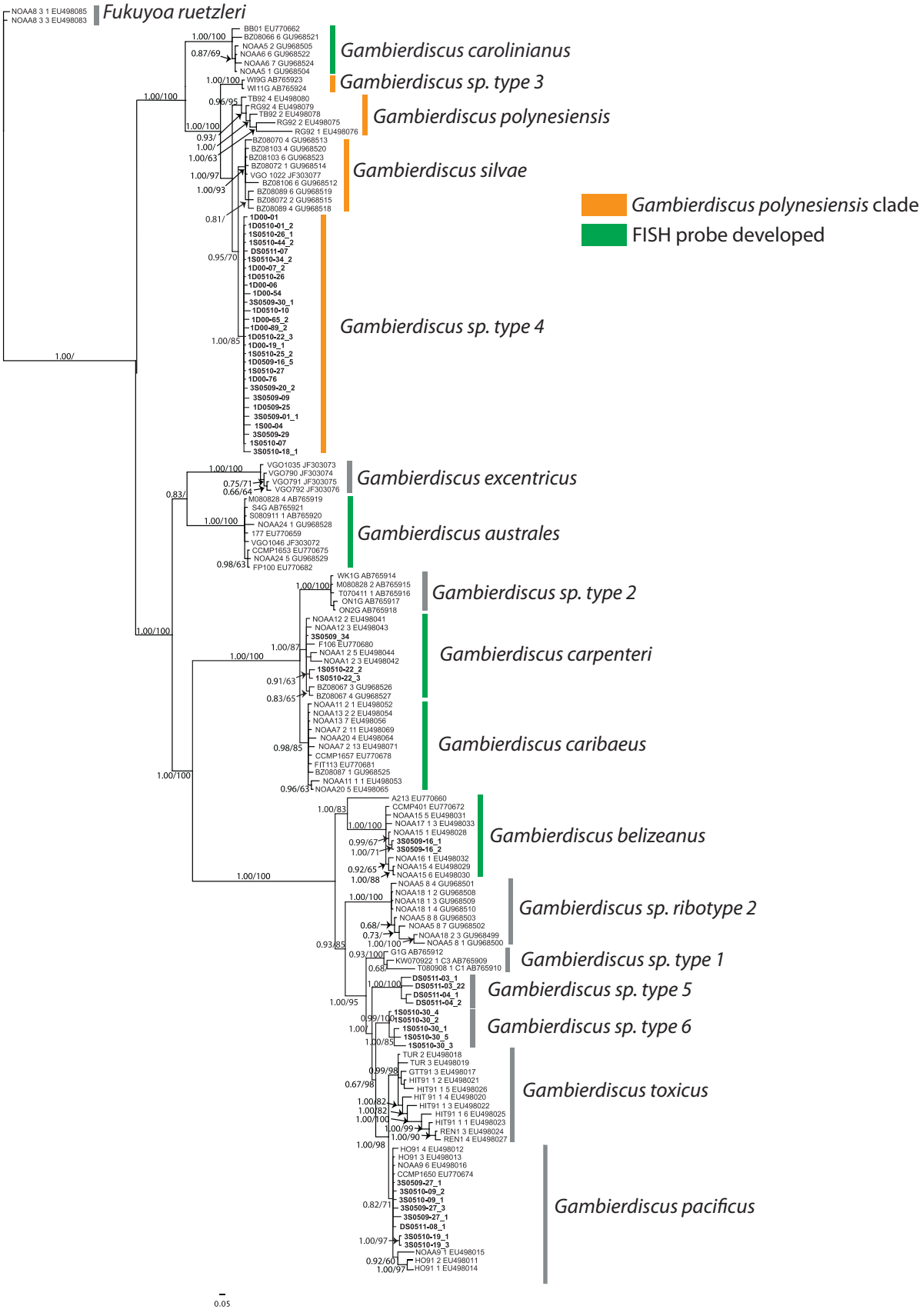
Ciguatera fish poisoning (CFP) is a syndrome caused by the bioaccumulation of lipophilic ciguatoxins in coral reef fish and their subsequent consumption by humans. These phycotoxins are produced by *Gambierdiscus*, a tropical epiphytic dinoflagellate genus that lives on many varieties of macroalgae, but also may occur on dead corals and in sand. Recently, the diversity with the *Gambierdiscus* genus has been revealed, with 11 species now identified, many of which co-occur. These species also differ significantly in toxicity. The ability to accurately and quickly distinguish *Gambierdiscus* species and determine community composition is central to assessing CFP risk, yet most *Gambierdiscus* species are indistinguishable using light microscopy. In order to investigate the dynamics of *Gambierdiscus* species diversity over time and among locations, new tools for species identification are needed. Here, fluorescence in situ hybridization (FISH) probes were designed to differentially label six species, permitting their enumeration using fluorescence microscopy. This technology enables the assessment of community composition and determination of cell abundances. Probes detecting *G. australes*, *G. caribaeus*, *G. carolineanus, G. carpenteri, G. belizeanus*, and the *G. polynesiensis* clade were designed using alignments of large subunit ribosomal sequences. They were tested on cultures of *Gambierdiscus,* and using field samples collected from the Florida Keys, Saint Thomas, USVI, and Hawai'i. A seasonal progression of samples over a six-month period was examined from the Florida Keys, showing shifts in *Gambierdiscus* community composition as overall *Gambierdiscus* cell abundance increased.

Introduction

Ciguatera fish poisoning (CFP) is a syndrome caused by the bioaccumulation of lipophilic ciguatoxins in coral reef fish and subsequent consumption by humans (Lehane & Lewis, 2000; Scheuer et al., 1967). These phycotoxins are produced by *Gambierdiscus*, a tropical epiphytic dinoflagellate genus that lives on many varieties of macroalgae but also may occur on dead corals and sand. Globally, tens of thousands of individuals are likely afflicted with ciguatera on an annual basis, with up to 10% of the local population on some islands in endemic areas becoming ill (Lora E. Fleming et al., 1998). CFP outbreaks have long been associated with ecological disturbances to reef environments, likely due to proliferation of *Gambierdiscus'* preferred substrates of algae and dead coral (Kohlerl & Kohler, 1992; Ruff, 1989). This presents an added concern that as the health of global coral reefs suffer due to damage from overfishing, pollution, and bleaching events, the threat of CFP will increase concurrently. *Gambierdiscus* species have been shown to have different temperature ranges of growth as well as different ranges in toxicity (Holland et al., 2013; Kibler et al., 2012). These physiological differences make understanding *Gambierdiscus* species community composition essential to calculating risk of CFP. However, *Gambierdiscus* species are almost indistinguishable under the light microscope. Current techniques to identify cells to the species level include sequencing of large subunit ribosomal DNA (LSU rDNA) sequences or scanning electron microscopy, techniques that are impractical to apply on a large scale. Another technique using semi-quantitative qPCR has also been developed for large scale analysis however, due to varying levels of rDNA within *Gambierdiscus* cells, this assay does not have the same quantitative strength of microscope counts (Vandersea et al., 2012). This chapter shows that the development of species-specific fluorescent *in situ* hybridization (FISH) probes allows for analysis of field samples to determine *Gambierdiscus* community composition while preserving the accuracy of counting cells under the light microscope. This differentiation of *Gambierdiscus* species allows the elaboration of community dynamics and shifting patterns of community toxicity and therefore can elucidate the changing risk of ciguatera fish poisoning and provide a needed tool for monitoring of CFP risk.

FISH probes have been used previously to target LSU rRNA in order to discriminate between dinoflagellate species (e.g. *Alexandrium* (Anderson, 1995; Scholin & Anderson, 1993)). In cases where the target is only a minor member of a phytoplankton assemblage or where there is a great degree of detritus in the samples, as is true in both cases for *Gambierdiscus*, FISH probes enable accurate identification of cells (Anderson, 1995). Probes detecting *G. australes*, *G. belizeanus, G. caribaeus*, *G. carolineanus, G. carpenteri*, and the *G. polynesiensis* clade (*G. polynesiensis* and *G. silvae)* were designed using alignments of large subunit ribosomal sequences in the D1-D3 and D8-D10 regions of LSU rRNA and tested using culture and field samples from the Florida Keys, Saint Thomas, USVI, and Hawai'i. These species are commonly isolated from sites in the Caribbean, Pacific Ocean, or both regions. From one site in the Florida Keys, a seasonal progression of samples over a six-month period from June to December 2013 is examined.

Figure 3.1 *Gambierdiscus* Phylogeny showing Probe Targets
Adapted from Xu et al., 2014 (Xu et al., 2014), phylogeny of *Gambierdiscus* species from D8-D10 region of the LSU rRNA gene. Scale bar = 0.05 substitutions per site. Support values are Bayesian posterior probability and bootstrap support values from maximum likelihood analysis. In orange is highlighted the '*G. polynesiensis* clade'. In green, species for which this study has developed species-specific probes are highlighted.


Fukuyoa ruetzleri
Gambierdiscus carolinianus
Gambierdiscus sp. type 3
Gambierdiscus polynesiensis
Gambierdiscus silvae
Gambierdiscus sp. type 4
Gambierdiscus excentricus
Gambierdiscus australes
Gambierdiscus sp. type 2
Gambierdiscus carpenteri
Gambierdiscus caribaeus
Gambierdiscus belizeanus
Gambierdiscus sp. ribotype 2
Gambierdiscus sp. type 1
Gambierdiscus sp. type 5
Gambierdiscus sp. type 6
Gambierdiscus toxicus
Gambierdiscus pacificus
Gambierdiscus polynesiensis clade
FISH probe developed
0.05

Methods

Probe Design and Testing

Fluorescent probes were designed to target the large subunit (LSU) region of rRNA of seven species of *Gambierdiscus: G. australes*, *G. belizeanus, G. caribaeus*, *G. carolineanus, G. carpenteri, G. polynesiensis* and *G. silvae*. Publically available sequences from the D1-D3 and D8-D10 domains of the LSU were aligned and regions of conservation chosen within species that contain at least two nucleotides of difference between species. Probes were conjugated to a fluorophore (Cy3), hybridized against formalin and methanol preserved samples, and analyzed under the fluorescent microscope (Zeiss M1 Imager) to determine fluorescence of cells. Probes were tested against cultures of *G. australes*, *G. belizeanus, G. caribaeus, G. carolinianus, G. carpenteri, G. pacificus, G. polynesiensis, G. silvae, G. toxicus*, and *G. ribotype II*. Cultures in separate hybridization reactions but within the same hybridization procedure were hybridized with a positive control universal probe that targets small subunit (SSU) rRNA (Univ-1390, Zheng et al., 1996) and also with no probe as a negative control. These reactions were used to determine both that cells' rRNA complement was sufficient and able to be labeled by FISH probe, and that probe fluorescence was dramatically different than cell autofluorescence. Probes that passed this initial testing were then conjugated with one of three fluorophores (Alexa Fluor® 488, 532, or 594) if used in triplicate or one of two fluorophores (Cy3 or Alexa Fluor® 488) if used in duplicate and tested against field samples in conjunction with one or two other probes.

Probe Specificity Testing

Probes were tested against cultures of *Gambierdiscus* to determine their species specificity. *Gambierdiscus* species included *G. australes* ('CCMP 1653'), *G. belizeanus* ('CCMP399', 'BP Apr11_7', 'BP Mar10_6'), *G. caribaeus* ('GTBNAC1', 'BP Aug08'), *G. carolineanus* ('BB Apr10_6', 'BP Mar10_1'), *G. carpenteri* ('HGB6'), *G. pacificus* ('350509_271', 'D50511-08'), *G. polynesiensis* ('RIK-8', 'RAI-1', 'RG-92'), *G. silvae* ('SH Apr11-1', 'TRL23', 'FC May10_9') and *G. ribotype II* ('BP Mar10_5').

Sample Fixation

Samples are fixed with 750ul of formalin in 14ml of culture or field sample, spun down, supernatant removed, and then replaced with methanol (volume depending on the density of the culture or field sample. Samples were then allowed to sit at -20°C for at least 24 hours prior to their labeling through whole cell hybridization.

Whole Cell Hybridization

As described in (Anderson et al., 2005), a whole cell (WC) hybridization method was employed to label cells with the FISH probes. A sample volume of 1ml of field sample, or 300ul of culture, was hybridized against FISH probes at a concentration of 2ng/ul within a hybridization manifold at 52, 53 or 55 degrees for an hour, depending on the probe mixture. During the wash step Calcofluor White was added to field samples to enable *Gambierdiscus* identification at the genus level due to its labeling of dinoflagellate thecal plates (20ul Calcofluor White in 15ml 0.2 SET wash solution). Hybridization filters were then placed on slides with glycerol and stored at -4°C until counted under the fluorescent microscope.

Field Sample Testing

Field samples from the Florida Keys and Saint Thomas are routinely collected as part of an ongoing study to monitor *Gambierdiscus* abundance (CiguaHAB). Monthly samples from sites are collected from macroalgae and artificial substrate (tiles).

Macroalgal samples from Saint Thomas, USVI, were used in preliminary testing. Samples with a high amount of detritus and no or very low numbers of *Gambierdiscus* cells were chosen as background. They were spiked with known *Gambierdiscus* species from culture and used to determine if probe concentrations were sufficient to label cells when they were a minor fraction of a larger benthic community. Probes were tested at 2ng/ul, 3ng/ul, and 4ng/ul of hybridization reaction mix. Since cells were easily visualized at 2ng/ul hybridization mix, this concentration was chosen for the rest of this analysis. However, also due to samples from Saint Thomas being commonly high in detritus, samples from the Florida Keys was chosen for an initial field study into the efficacy of these probes as they are often much easier to enumerate.

Samples from a site in the Florida Keys, Heine Grassbed (HGB), were selected for a seasonal comparison of *Gambierdiscus* community diversity (Table 3.1). Samples spanned seven months of collection in 2013 and from macroalgae genera including *Thalassia* and *Halimeda*. These samples were labeled with individual probes detecting *G. belizeanus, G. caribaeus*, *G. carolineanus, G. carpenteri*, and the *G. polynesiensis* clade attached to the Cy3 fluorophore. This analysis was completed prior to combining probes with multiple fluorophores.

Field samples were also collected from Wai'Ōpae Tide Pools, Hawai'i as part of a study of temperature effects on *Gambierdiscus* community composition (Chapter 4). These samples were largely labeled with combinations of three and two probes as described in the Pacific assay (Table 3.2), though some were labeled with singular Cy3 attached probes. They illustrate the efficacy of the Pacific probe assay.

Date	Site	cells/ml fixed	Genus	Algae
6/12/13	HGB	44	*Thalassia*	Dead *Thalassia*
6/12/13	HGB	41	*Halimeda*	*Halimeda incrassata*
7/9/13	HGB	40	*Thalassia*	*Thalassia testudinum*
7/9/13	HGB	38	*Halimeda*	*Halimeda incrassata*
8/14/13	HGB	71	*Halimeda*	*Halimeda incrassata*
8/14/13	HGB	38	*Halimeda*	*Halimeda incrassata*
9/11/13	HGB	61	*Thalassia*	*Thalassia testudinum*
9/11/13	HGB	46	*Thalassia*	Dead *Thalassia*
10/10/13	HGB	107	*Halimeda*	*Halimeda incrassata*
10/10/13	HGB	79	*Halimeda*	*Halimeda incrassata*
11/11/13	HGB	379	*Halimeda*	*Halimeda incrassata*
11/11/13	HGB	202	*Halimeda*	*Halimeda incrassata*
12/16/13	HGB	372	*Halimeda*	*Halimeda incrassata*

Table 3.1 HGB 2013 Samples
Samples from the Heine Grassbed (HGB) site in the Florida Keys listing their total abundance of *Gambierdiscus* cells and genera of macroalgae they were harvested from.

Probe Combinations
Probes were combined based on their hybridization temperature (Tm) ranges and presence of species in geographic locations. Two groups of probes were designed for the Pacific and the Caribbean (Table 3.2, Table 3.3). Complementary fluorophores were hybridized to probes within groups allowing enumeration of multiple species from the same sample (Figure 3.2).

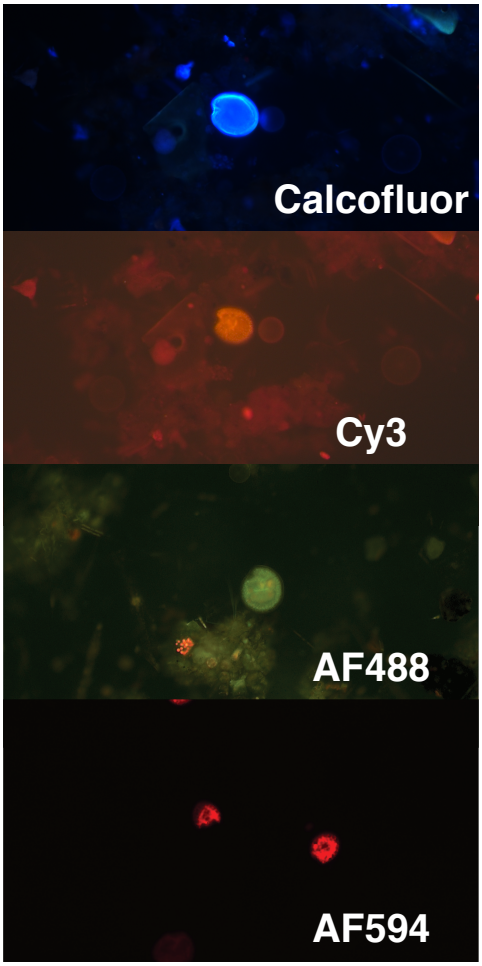


Figure 3.2 Fluorophore and Calcofluor Labeling of *Gambierdiscus* Cells
Figure shows different *Gambierdiscus* cells labeled by probes hybridized to Cy3, AF488, and AF594 under Cy3, FITC, and TxRd filters respectively. In the top panel, The same cell labeled by a Cy3-hybridized probe is shown under a DAPI filter with Calcofluor White labeling of its thecal plates. This permits identification at the genus level.

Caribbean						
Group	**Probe**	**Target Species**	**Sequence**	**Tm**	**Filter**	**Fluorophore**
1	Gcarib_D1D3	*G. caribaeus*	TGAGACCCACATGTGGAGATTC	**53**	FITC	AF488
	Gbeliz_D8D10	*G. belizeanus*	AGATCAGTACGCCAGAGTGACTA		Cy3	Cy3
2	Gcarpent_D1D3	*G. carpenteri*	TGATGTAACGCAAGACGCACAG	**55**	FITC	AF488
	Gribo1_D8D10	*G. polynesiensis clade*	CGATCAGAGACATACTTTGGCGC		Cy3	AF532
	Gcarol_D8D10	*G. carolineanus*	AGCAAGCCACAGATCCACTGAG		TxRd	AF594

Table 3.2: Caribbean and Gulf of Mexico Probe Combinations
FISH probes in Group 1 or 2 were used simultaneously to label Caribbean and Gulf of Mexico samples. Table shows probe target species, sequence, melting temperature (T_m), microscope filter, and fluorophore used for illumination.

Pacific						
Group	**Probe**	**Target Species**	**Sequence**	**Tm**	**Filter**	**Fluorophore**
1	Gcarib_D1D3	*G. caribaeus*	TGAGACCCACATGTGGAGATTC	**52**	FITC	AF488
	Gaust_D1D3	*G. australes*	TGCCAATCCAGTTGTGTATCTC		Cy3	Cy3
2	Gcarpent_D1D3	*G. carpenteri*	TGATGTAACGCAAGACGCACAG	**53**	FITC	AF488
	Gbeliz_D8D10	*G. belizeanus*	AGATCAGTACGCCAGAGTGACTA		Cy3	AF532
	Gpoly_D8D10	*G. polynesiensis clade*	CTCCGCCAGTGACGTTAAGTAG		TxRd	AF594

Table 3.3 Pacific Ocean Probe Combinations
FISH probes in Group 1 or 2 were used simultaneously to label Pacific samples. Table shows probe target species, sequence, melting temperature (T_m), microscope filter, and fluorophore used for illumination.

Results

Probe Culture Testing

Probes were tested against cultures of *Gambierdiscus* to determine their specificity (Table 3.4, 3.5, 3.6). *Gambierdiscus* species included *G. australes* ('CCMP 1653'), *G. belizeanus* ('CCMP399', 'BP Apr11_7', 'BP Mar10_6'), *G. caribaeus* ('GTBNAC1', 'BP Aug08'), *G. carolineanus* ('BB Apr10_6', 'BP Mar10_1'), *G. carpenteri* ('HGB6'), *G. pacificus* ('350509_271', 'D50511-08'), *G. polynesiensis* ('RIK-8', 'RAI-1', 'RG-92'), *G. silvae* ('SH Apr11-1', 'TRL23', 'FC May10_9') and *G. ribotype II* ('BP Mar10_5'). Probes Gaust_D1D3, Gcarib_D1D3, Gbeliz_D8D10, Gcarpent_D1D3, and Gcarol_D8D10 only detected their target species but probes Gpoly_D8D10 and Gribo1_D8D10 both detected *G. polynesiensis* and *G. silvae*. For all known ribosomal large subunit sequences both Gpoly_D8D10 and Gribo1_D8D10 have at least two base pairs of difference between the species, however, these species are very closely related (Figure 3.1). Therefore, for this and future analysis both the probes Gpoly_D8D10 and Gribo1_D8D10 have been redefined to broadly detect species within the '*G. polynesiensis* clade' (Figure 3.1).

			Culture					
			G. australes	*G. belizeanus*			*G. caribaeus*	
Probe	Target Species	Tm °C	CCMP 1653	CCMP 399	BP Apr 11_7	BP Mar 10_6	GTBNAC1	BP Aug08
Gaust_D1D3	*G. australes*	52	**Positive**	Negative	Negative	Negative	Negative	Negative
Gcarib_D1D3	*G. caribaeus*	52/53	Negative	Negative	Negative	Negative	**Positive**	**Positive**
Gbeliz_D8D10	*G. belizeanus*	53	Negative	**Positive**	**Positive**	**Positive**	Negative	Negative
Gpoly_D8D10	*G. polynesiensis*	53	Negative	Negative	Negative	Negative	Negative	Negative
Gcarpent_D1D3	*G. carpenteri*	53/55	Negative	Negative	Negative	Negative	Negative	Negative
Gcarol_D8D10	*G. carolinianus*	55	Negative	Negative	Negative	Negative	Negative	Negative
Gribo1_D8D10	*G. polynesiensis clade*	55	Negative	Negative	Negative	Negative	Negative	Negative

			Culture					
			G. carolinianus		*G. carpenteri*	*G. polynesiensis*		
Probe	Target Species	Tm °C	BB Apr10_6	BP May 10_1	HGB6	RIK-8	RAI-1	RG-92
Gaust_D1D3	*G. australes*	52	Negative	Negative	Negative	Negative	Negative	Negative
Gcarib_D1D3	*G. caribaeus*	52/53	Negative	Negative	Negative	Negative	Negative	Negative
Gbeliz_D8D10	*G. belizeanus*	53	Negative	Negative	Negative	Negative	Negative	Negative
Gpoly_D8D10	*G. polynesiensis*	53	Negative	Negative	Negative	**Positive**	**Positive**	**Positive**
Gcarpent_D1D3	*G. carpenteri*	53/55	Negative	Negative	**Positive**	Negative	Negative	Negative
Gcarol_D8D10	*G. carolinianus*	55	**Positive**	**Positive**	Negative	Negative	Negative	Negative
Gribo1_D8D10	*G. polynesiensis clade*	55	Negative	Negative	Negative	**Positive**	**Positive**	**Positive**

			Culture					
			G. pacificus		*G. silvae*			*G. ribotype II*
Probe	Target Species	Tm °C	350509_271	D50511-08	SH Apr11-1	TRL23	FC May 10_9	BP Mar 10_5
Gaust_D1D3	*G. australes*	52	Negative	Negative	Negative	Negative	Negative	Negative
Gcarib_D1D3	*G. caribaeus*	52/53	Negative	Negative	Negative	Negative	Negative	Negative
Gbeliz_D8D10	*G. belizeanus*	53	Negative	Negative	Negative	Negative	Negative	Negative
Gpoly_D8D10	*G. polynesiensis*	53	Negative	Negative	**Positive**	**Positive**	**Positive**	Negative
Gcarpent_D1D3	*G. carpenteri*	53/55	Negative	Negative	Negative	Negative	Negative	Negative
Gcarol_D8D10	*G. carolinianus*	55	Negative	Negative	Negative	Negative	Negative	Negative
Gribo1_D8D10	*G. polynesiensis clade*	55	Negative	Negative	**Positive**	**Positive**	**Positive**	Negative

Table 3.4, 3.5, 3.6: Probe testing results against *Gambierdiscus* cultures
Cultures were preserved in formalin and methanol and hybridized against a probe with a Cy3 fluorophore. Probes that detected a culture and were highly visible under the fluorescent microscope are denoted 'Positive'. Probes that failed to detect a culture are denoted 'Negative'.

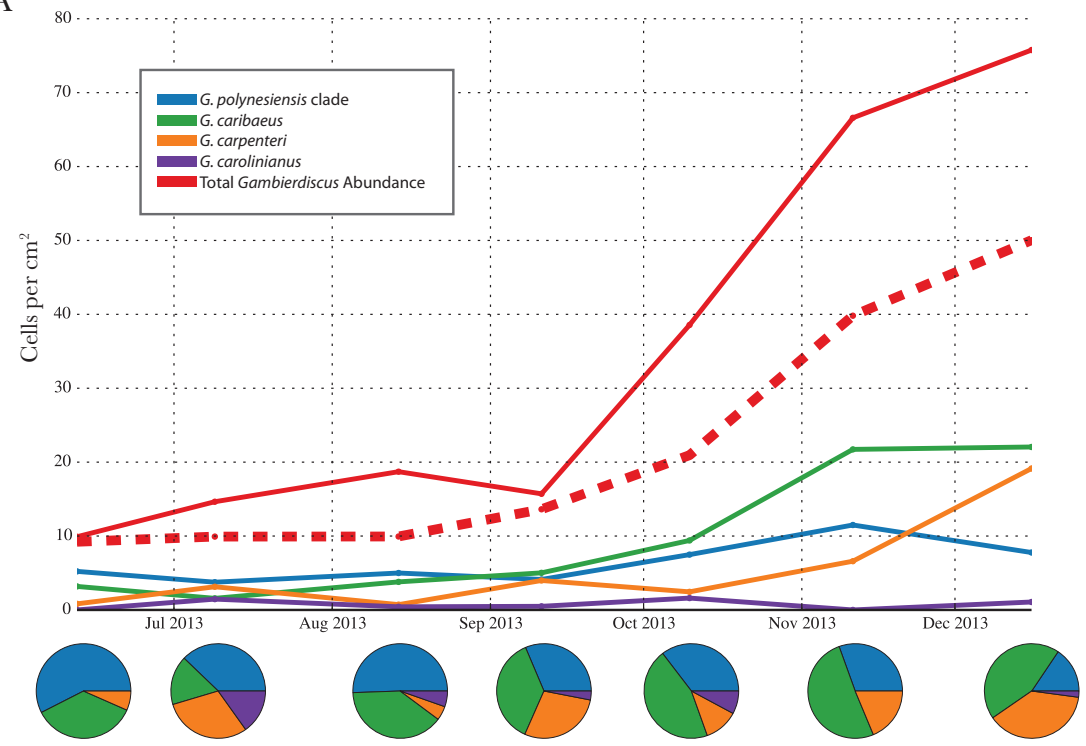
Probe Field Sample Analysis

Samples from Heine Grassbed (HGB) in 2013 were labeled with individual probes detecting *G. belizeanus, G. caribaeus*, *G. carolineanus, G. carpenteri*, and the *G. polynesiensis* clade. Shifts in *Gambierdiscus* community structure over time were apparent with a decrease in relative abundance of the *G. polynesiensis* clade and an increase in relative abundance of *G. caribaeus* and *G. carpenteri* as *Gambierdiscus* abundance overall increased at this site from June to December 2013 (Figure 3.3).

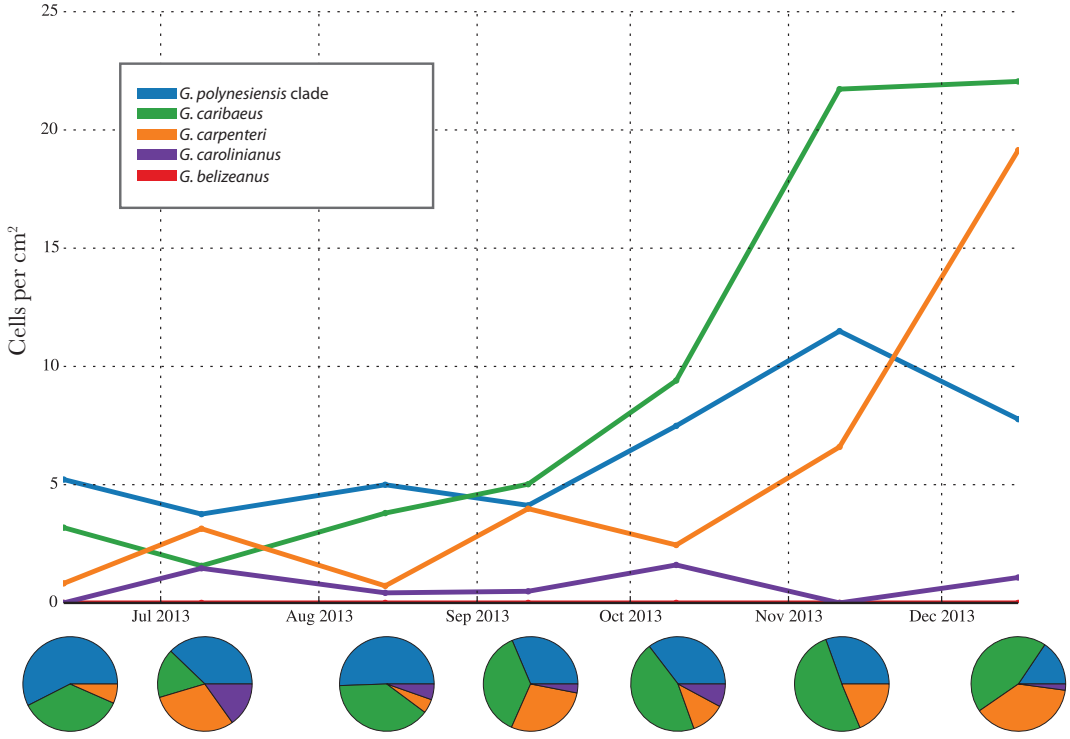
Field samples from Wai'Ōpae Tide Pools, Hawai'i, were labeled with probes detecting *G australes, G. belizeanus, G. caribaeus*, *G. carpenteri*, and the *G. polynesiensis* clade as described in Table 3.2. All species were found present at the site. Probes worked in combinations of three fluorophores to detect up to three species at once. Ecological significance of species composition data is discussed in Chapter 4 but visual images of labeled cells are presented here to show probe efficacy (Figure 3.4).

Figure 3.3 HGB Gambierdiscus Community Composition and Abundance
Using FISH probes, species abundance was determined and averaged over each month. (A) Total abundance of *Gambierdiscus* in samples through genus level counts is shown in a red line and summed abundance of species is shown in a dotted red line. (B) *Gambierdiscus* species abundance over the sampling period without total abundance illustrating changes in relative abundance. Abundance is in cells per cm^2 macroalgae surface area calculated from algal blade or thallus area.

A


G. polynesiensis clade
G. caribaeus
G. carpenteri
G. carolinianus
Total Gambierdiscus Abundance
Cells per cm2
80
70
60
50
40
30
20
10
0
Jul 2013
Aug 2013
Sep 2013
Oct 2013
Nov 2013
Dec 2013

B


G. polynesiensis clade
G. caribaeus
G. carpenteri
G. carolinianus
G. belizeanus
Cells per cm2
25
20
15
10
5
0
Jul 2013
Aug 2013
Sep 2013
Oct 2013
Nov 2013
Dec 2013

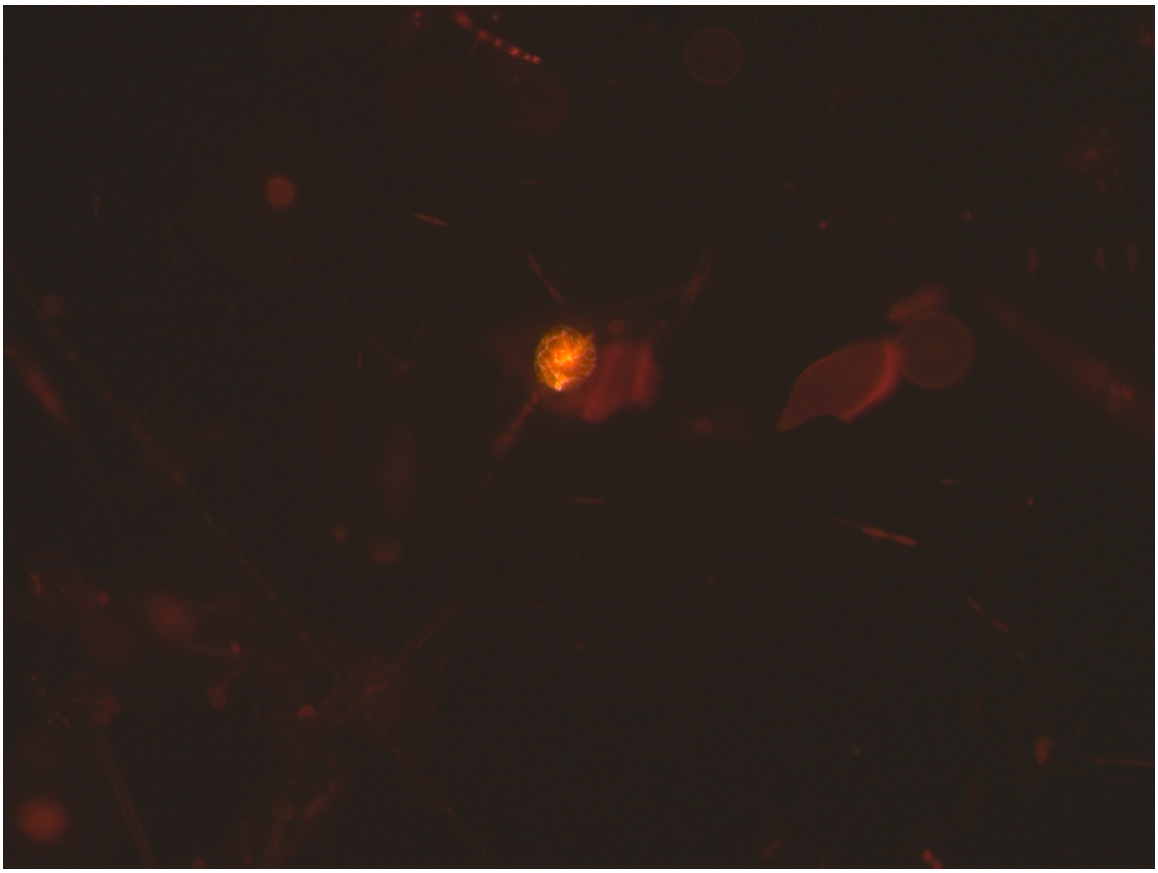


Figure 3.4 Labeled Cell in Field Sample
Gambierdiscus belizeanus labeled with probe attached to Cy3 fluorophore in a field sample from Hawai'i.

Discussion

The development of species-specific fluorescent *in situ* hybridization (FISH) probes allows analysis of field samples to determine *Gambierdiscus* community composition. Due to *Gambierdiscus* species' cryptic diversity and variance in toxicity, species-specific enumeration is an important component in determining regional risk of ciguatera fish poisoning. Other methods previous to this study require prohibitive time and effort for routine monitoring (sequencing or electron micrograph imaging) or are only semi-quantitative (qPCR). This study found that novel FISH probes were able to elucidate seasonal dynamics of a *Gambierdiscus* community at Heine Grassbed (HGB), Florida Keys. At this site as *Gambierdiscus* overall abundance increased from June to December 2013, the relative abundance of *G. caribaeus* and *G. carpenteri* increased while that of the *G. polynesiensis* clade decreased. This showed a seasonal progression of *Gambierdiscus* species, with some species favored in earlier months, while others became more abundant as the bloom progressed. Since different species may contain different levels of toxicity, determining which species are favored under bloom conditions aids in determining risk of CFP for this area.

Furthermore, species composition may have a greater affect on risk of CFP than total *Gambierdiscus* abundance. For example, if an increase in *Gambierdiscus* abundance was due largely to a nontoxic species, it would not present any additional risk. In these field samples, there was also a discrepancy between the total abundance of *Gambierdiscus* and the summed abundance of all species detected (Figure 3.3A). This could be due to the presence of another species for which there is currently no FISH probe developed.

Comparison of Methods and Proposed Monitoring Techniques

In addition to this novel FISH probe identification of *Gambierdiscus* species, there is another molecular method to differentiate the *Gambierdiscus* community, a semi-quantitative qPCR assay (Vandersea et al., 2012). A combination of approaches to *Gambierdiscus* community determination could be used, as suggested in Kibber et al., 2014 (Proceedings of the 15th ICHHA). The qPCR assay is able to detect species at low cell numbers that may be missed if a cell does not happen to be contained within the aliquot of sample used for FISH labeling, however, qPCR limits of detection should also be considered. If too high a volume of a field sample is combined, *Gambierdiscus* cells may become too small of a constituent for qPCR detection. Through DNA extraction, the qPCR assay can theoretically look at a greater volume of sample more rapidly than FISH analysis depending on the assay's sensitivity. In this way, samples could be examined for the presence or absence of an array of species and then FISH probes applied to enumerate a chosen selection of species that may be important for analysis due to their toxicity. Furthermore, in well characterized sites where only a few species are known to be toxic, their abundances could be monitored through FISH at low cost. Total *Gambierdiscus* abundance could be measured at the same time through Calcofluor White labeling of the hybridizations. At this time there also exists the ability within FISH probes to detect species in the *G. polynesiensis* clade, while no qPCR assay for these species has been made. Knowledge of new *Gambierdiscus* species is currently expanding (Fraga & Rodríguez, 2014) and discovery of new species will force reevaluation of current detection techniques. However, once it is discovered which groups of species of *Gambierdiscus* are the most toxic, and responsible for the highest risk of ciguatera, probes can be targeted towards these clades, encompassing a degree of genetic variability.

Conclusion

The development of species-specific fluorescent *in situ* hybridization (FISH) probes allows analysis of field samples to determine *Gambierdiscus* community composition. Due to *Gambierdiscus* species' cryptic diversity and variance in toxicity, species-specific enumeration is an important component in determining regional risk of ciguatera fish poisoning. This novel approach allowed the quantitative assessment of species along a six-month seasonal increase in *Gambierdiscus* abundance in Heine Grassbed (HGB), Florida Keys. A shift in *Gambierdiscus* community composition occurred as total cell abundance increased, with potential implications for risk of ciguatera fish poisoning at this site. Pacific Ocean probe combinations were also shown to enumerate species in Hawai'i. By analyzing *Gambierdiscus* community composition across multiple environments, differences in species abundance can be linked to environmental parameters improving our ability to calculate current and changing risk of CFP worldwide.

Funding Sources:
USFDA (F223201000060C)
NOAA NOS (CiguaHAB; NA11NOS4780060 and NA11NOS4780028);
COHH - Woods Hole Center for Oceans and Human Health – NSF Grant OCE-1314642 and NIEHS 1P01-ES021923-01
WHOI Ocean Ventures Fund; WHOI Coastal Institute Research Proposal Award

Chapter 4

Gambierdiscus Population Structure across a Thermally Variable Environment: Wai'Ōpae Tide Pools, Hawai'i

Abstract

Species within the *Gambierdiscus* genus produce ciguatoxin precursors that bioaccumulate and are biotransformed into lipophilic ciguatoxins in coral reef fish, and if consumed by humans, these toxins can cause ciguatera fish poisoning (CFP), or 'ciguatera'. Globally, tens of thousands of individuals are likely afflicted with CFP on an annual basis causing significant public health and economic concerns. Prior studies have shown that multiple *Gambierdiscus* species co-occur, and have significant differences in toxicity and temperature tolerance. Prolonged periods of elevated water temperatures such as those associated with global warming are hypothesized to result in increased *Gambierdiscus* cell densities and in turn, increased flux of ciguatoxins into the food chain. The ability to accurately distinguish *Gambierdiscus* species and determine community composition is therefore central to assessing CFP risk. Furthermore, due to changing water temperatures, species' ranges may expand or shift, changing future *Gambierdiscus* community composition and areas that experience ciguatera. By analyzing *Gambierdiscus* communities in areas that experience large temperature variability and that are known to host multiple *Gambierdiscus* species, the potential effects of changing water temperatures on species' ranges can be explored. Here, fluorescence in situ hybridization (FISH) probes were used to differentially label five *Gambierdiscus* species in Wai'Ōpae Tide Pools, Hawai'i, where a series of interconnected tide pools exhibit differences in temperature and allow us to examine the effects of this variation on *Gambierdiscus* species diversity across seasonal and spatial thermal gradients. Relative abundance of each species was determined using cell counts from short-term deployments of artificial substrates (window screens), and from coral rubble, which reflects longer-term colonization and growth. Community composition was compared across pools and also seasonally, and abundance data were used to model *Gambierdiscus* species composition across temperature regimes. Growth rates of strains isolated from field samples were also measured using a temperature gradient bar to determine their physiological tolerance for extreme temperatures. Field data revealed that all five species of *Gambierdiscus* were present. *Gambierdiscus australes* and the *G. polynesiensis* clade were the most abundant. *Gambierdiscus belizeanus*, *G. carpenteri*, and *G. polynesiensis* clade had higher relative abundance at lower temperatures while *G. caribaeus* and the unlabeled fraction of cells had higher relative abundance at higher seasonal mean temperatures. *Gambierdiscus australes* had high relative abundance across all modeled temperatures. Temperature growth curves for two *G. australes* strains and one *G. caribaeus* strain support this result as *G. caribaeus* has higher growth at higher temperatures while *G. australes* exhibited a wider range of thermal tolerance. Combined with other laboratory-derived growth rates in previous studies, the results from this study illustrate how shifts in temperature dramatically affect *Gambierdiscus* community composition. This represents the first time that changes in *Gambierdiscus* community composition have been resolved in a comparative field study. These results also show the site-specific nature of local *Gambierdiscus* diversity and physiology as even within a small geographic area, community composition varied widely. Importantly, experiencing periodic higher temperatures locally did not translate to higher *Gambierdiscus* abundances, but instead led to a shift in species composition and lower total cell abundance. Thus temperature variability may have a different effect on *Gambierdiscus* species growth than changes in mean temperature alone. Overall, this study found complex dynamics at play between environmental conditions and *Gambierdiscus* community composition that illustrate the difficulty of predicting future risk of ciguatera fish poisoning amid climate change conditions.

Introduction

Ciguatera fish poisoning (CFP), or 'ciguatera', is a syndrome where phycotoxins produced by species within the *Gambierdiscus* genus bioaccumulate and are biotransformed into lipophilic ciguatoxins in coral reef fish. When consumed by humans, these toxins can cause severe neurological and gastrointestinal symptoms (Lehane & Lewis, 2000; Scheuer et al., 1967). Globally CFP affects more people than any other seafood illness caused by toxic algae, and was reported to be to be the most common marine biotoxin responsible for food poisoning worldwide (EFSA, 2010; L.E. Fleming et al., 2006). CFP outbreaks have long been associated with ecological disturbances to reef environments, likely due to proliferation of *Gambierdiscus'* preferred substrates of algae and dead coral (Kohlerl & Kohler, 1992; Ruff, 1989). This presents an added concern that as the health of global coral reefs suffer due to damage from overfishing, pollution, and bleaching events, the threat of CFP will increase concurrently. Due to its responsiveness to environmental disturbance, CFP has been described as a sensitive indicator of changing conditions in tropical marine ecosystems (Hales et al., 1999). Several studies have predicted that as global sea surface temperatures increase, *Gambierdiscus* will expand its geographic range, increasing the number of people at risk of ciguatera fish poisoning (Dickey & Plakas, 2010; P. A. Tester et al., 2010; Villareal et al., 2007). Since no method is currently available to rapidly and cheaply test fish for ciguatoxins, fishermen instead use known poisoning histories of localized fishing areas to determine the risk of their catch. These histories may lose accuracy as climate-driven warming influences the range and abundance of *Gambierdiscus* species, emphasizing the need to better understand how increasing seawater temperatures and concomitant changes to reef communities influence *Gambierdiscus* population structure and physiology (P. A. Tester et al., 2010; Tosteson, 2004).

Prior studies have shown that multiple *Gambierdiscus* species co-occur, and have significant differences in toxicity and optimum temperature growth ranges (Holland et al., 2013; Kibler et al., 2012; Xu et al., 2016). However, most *Gambierdiscus* species are indistinguishable under the light microscope. The ability to accurately identify *Gambierdiscus* species and determine community composition is therefore central to assessing CFP risk. Current techniques to identify cells to the species level include sequencing of large subunit ribosomal DNA (LSU rDNA) sequences or scanning electron microscopy, techniques that are impractical to apply on a large scale. Another technique using semi-quantitative qPCR has also been developed for large scale analysis however, due to varying levels of rDNA within *Gambierdiscus* cells, this assay does not have the same quantitative strength of microscope counts (Vandersea et al., 2012). Novel fluorescence in situ hybridization (FISH) probes (Chapter 3) were developed and used to differentially label five species, thus permitting their identification and abundance in field samples. This technology enables accurate cell-based assessment of community composition. These probes were applied to samples from Wai'Ōpae Tide Pools, Hawai'i, a series of tide pools that exhibit differences in temperature, and permit investigation of the effects of natural temperature variations on *Gambierdiscus* species diversity.

In Wai'Ōpae, thermally variable but connected pools were originally formed from basaltic lava. The region remains heavily influenced by nearby geothermic activity through an influx of heated groundwater along the coast. Several pools within the back reef have similar mean water temperatures but experience daily fluctuations in temperature. Examining *Gambierdiscus* physiology and community dynamics in a region of highly variable temperature is a novel

approach to explore how *Gambierdiscus* population structure and geographic range might change with rising seawater temperatures. Furthermore, since the pools of Wai'Ōpae are not thermally invariable, but are instead cycling to different extents, they reveal how *Gambierdiscus* adapts to more or less thermally variable conditions. Seasonal comparisons between *Gambierdiscus* populations can thus illustrate differences in mean temperature tolerance, while comparisons among pools show the effects of short-term temperature variability.

Other environmental parameters such as salinity, nutrient concentrations, and coral cover also vary over Wai'Ōpae. Freshwater inputs close to shore cause lower salinity values primarily in surface layers. When water was sampled at depth of *Gambierdiscus* growth and natural substrates, salinity values were more constant across sites though there was still moderate evidence of lower salinities closer to shore at low tide. However, previous laboratory data have shown many *Gambierdiscus* species to be relatively tolerant of large ranges of salinity values (Xu et al., 2016). Nutrient concentrations also increase closer to shore. In another study of surface waters near these sampling sites, total dissolved nitrogen (TDN) ranged from 18 to 11 μmol L^{-1} from 0 to 100m from shore (Wiegner, Mokiao-Lee, & Johnson, 2016). Sites further from shore also have higher coral cover though corals are present near all sites. All these variables could also have an influence on *Gambierdiscus* community composition and cause variation between near-shore and more seaward pools though they were not extensively measured in this study.

To examine *Gambierdiscus* community composition, species relative abundances were determined using cell counts from short-term deployments of artificial substrates (window screens), and from coral rubble, the latter more indicative of longer-term colonization and growth of cells on a natural substrate. Species relative abundances were modeled across temperature regimes and community composition compared across pools and seasonally through hierarchical clustering and MDS plots. Growth experiments using a temperature gradient bar were then used to assess growth of field-isolated strains over a range of constant temperatures. Together, these data reveal a temperature-driven partitioning of *Gambierdiscus* species that has major implications with respect to climate change and potential ciguatera toxicity as coastal waters warm worldwide.

Methods

Field Site: Wai'Ōpae Tide Pools, Hawai'i
19.4877121°N, -154.8215246°W

Site 1 Top Lagoon
N19° 29.291' W154° 49.325'
Site 2 Bottom Lagoon
N19° 29.245' W154° 49.270'
Site 3 Control 2
N19° 29.281' W154° 49.153'
Site 4 Middle Lagoon
N19° 29.270' W154° 49.305'
Site 5 Control 1
N19° 29.210' W154° 49.225'
Site 6 Hot Pool
N19° 29.334' W154° 49.181'

Figure 4.1 Field Site Wai'Ōpae Tide Pools, Hawai'i
Figure shows the location of Wai'Ōpae Tide Pools, on the Island of Hawai'i. Highlighted shapes in orange are pool locations, stars denote sampling locations. Images adapted from Google Earth.

Wai'Ōpae Tide Pools is a Marine Life Conservation District on the southeastern shore of the Island of Hawai'i, in the Puna district about 30km southeast of Hilo (Figure 4.1). Wai'Ōpae consists of a series of pools among a rocky reef flat initially formed by a lava flow. Wai'Ōpae is located nearby the shield volcano, Kīlauea, which is the most active of five volcanoes that form

the Island of Hawai'i. The surrounding shoreline is influenced by geothermally-heated groundwaters entering the ocean. Wai'Ōpae is known for its rich coral growth and high coral diversity. It is ecologically important as a nursery for fishes and is under threat from a nearby private subdivision due to recreational use and potential household waste pollution.

Lately the corals at Wai'Ōpae have undergone several stress-inducing events. In August of 2014 Wai'Ōpae was particularly hard hit by hurricane Iselle and in 2015 Hawai'i as a whole suffered record high sea surface temperatures (Figure 4.3). These events may be favorable for *Gambierdiscus* abundance since *Gambierdiscus* has been shown to thrive on coral reef disruption. Furthermore, hurricane induced leakage of septic tanks from the nearby housing development could have introduced nutrients into the water, stimulating growth of macroalgae. Data from select pools, including the lagoon, show an increase in turf algal abundance following the hurricane from September to October 2014, from less than 20% to greater than 40% benthic cover with a decrease in live coral (M. Takabayashi, 'http://hilo.hawaii.edu/blog/misakita/category/coral-reef/'). Sampling for this study commenced in January 2015, five months after hurricane Iselle hit the region, and continued through the following summer, which was marked by record high seawater temperatures.

Temperature Measurements at Wai'Ōpae

Water temperatures in Wai'Ōpae, Hawai'i, are influenced from influxes of geothermally-heated groundwater from the shore and from the seaward side by the ocean. Daytime solar heating of tide pools and the cooling of pools at night also contribute to temperature variability. At high tide the pools are connected and there is a high degree of water circulation. This is in part driven by northeastern tradewind-generated swell. Temperature loggers were placed across the study site from January 2015 to January 2016. Due to the nature of the site as a popular tourist destination, several temperature loggers went missing during this study and had to be periodically replaced. This has resulted in gaps in the temperature record that were partially filled using nearby temperature data from another study (Dr. Misaki Takabayashi (UH Hilo); Site 3 replaced with 'Flag Pole' site, N19.489041, W154.818405) and by merging sites together: after the loss of the logger at Site 2, temperature data from Site 4 was used since both sites are in the same large pool (lagoon), relatively close together (57m apart), and temperature records from January showed them to be under very similar temperature regimes.

Salinity Gradient

The groundwater input at Wai'Ōpae creates a visible fresher mixed layer near sites closer to shore. Salinity measurements were taken at the depth of artificial substrate deployment to test benthic salinity that could affect *Gambierdiscus* growth. Salinity measurements at two time points (4/1/15 and 7/16/15) were collected at depth with whirl pack bags and measured by refractometer. Artificial substrate and rubble measurements were collected at depths whose salinity was less influenced by freshwater.

Field Sampling

To assess diversity and physiology of *Gambierdiscus* spp. at the study site, artificial substrate and rubble samples were collected from six lagoon and back reef pool locations and processed

following established methods (Lobel, Anderson, & Durand-Clement, 1988; P. A. Tester et al., 2014). Artificial substrates were constructed using black window screen, fishing line, fishing swivels to allow free movement, and a small float to suspend the screen in the water column (Figure 4.2) (P. A. Tester et al., 2014). Screens were 15.5 by 10.5cm. Volcanic basalt rocks were used as weights and wired to the bottom of the artificial substrate in the field. There was an abundance of volcanic rubble particular to this location that could be used for this purpose. As in previous studies, screen surface area for *Gambierdiscus* abundance measurements was calculated by considering the screen to be composed of cylindrical filaments resulting in a total surface area of 156.74cm^2 (P. A. Tester et al., 2014).

Artificial substrates were deployed for 24 and 72 hours at a time. Each individual screen was only used for one deployment while the rest of the apparatus was recycled for multiple uses. In January, March, and July six artificial substrates were deployed at each site for *Gambierdiscus* community analysis, three left for 24 hours and three left for 72 hours. Screens were collected underwater by carefully placing a large sealable plastic bag around the screen, detaching it from the rest of the apparatus, and then sealing the screen and surrounding water inside the bag. Bag contents were then transported back to shore for processing. In March three additional artificial substrates were placed at each site and then used for live culture generation.

To collect rubble samples, small rocks with turf algal cover were selected from the pool bottom and placed carefully into sealable plastic bags. Similar to artificial substrates, these samples were then brought back to shore for processing.

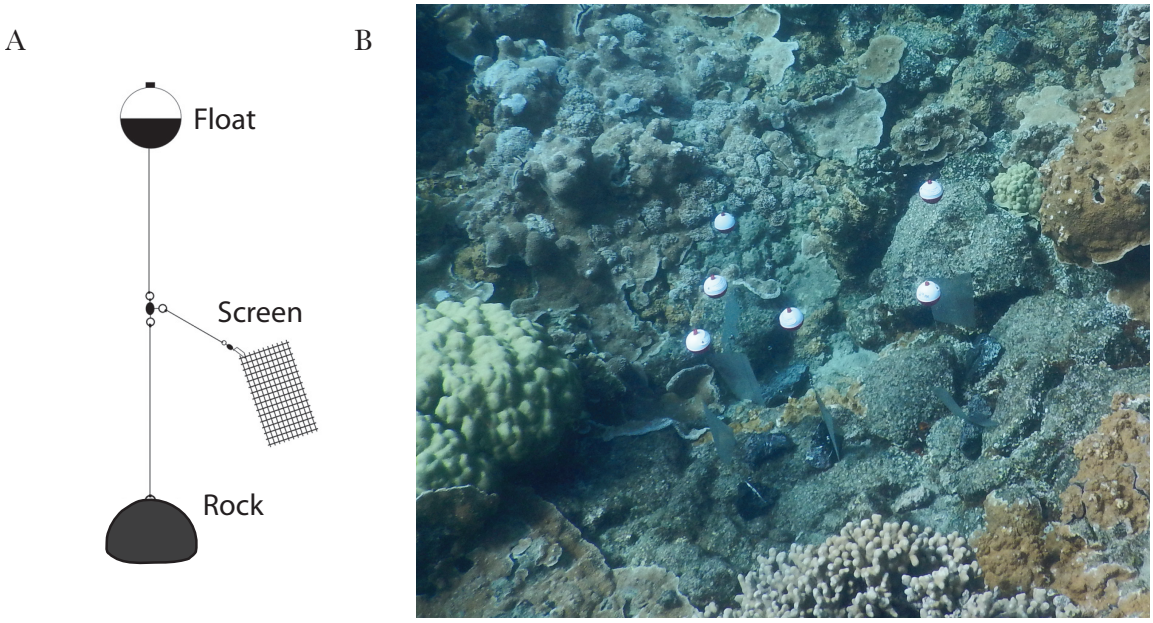


Figure 4.2: Artificial Substrate Sampling Device

Figure shows screen sampling device used in field study: (A) diagram adapted from (P. A. Tester et al., 2014) and (B) image of constructed devices deployed at Wai'Ōpae. A small fishing float was attached through a swivel with fishing line to a screen and a rock used as a weight. Sampling devices were replaced on location after sample collection and processing (below).

Sample Processing

For processing, field samples were shaken vigorously and their contents poured through a 100µm or 200µm, and 20µm sieves. Filtered seawater was then poured into the plastic bag and this process was repeated three times. The fraction of sample collected on the 20µm sieve was then backwashed into a 15ml falcon tube using filtered seawater and placed in a cooler for same-day transport to the University of Hawaii, Hilo, where samples were fixed with formalin (700ul of 37% formaldehyde into 14ml of field sample), centrifuged, and the supernatant removed and replaced with methanol (as described in Chapter 3). Samples were stored at -20°C until analysis. At the end of the sampling period, samples were shipped on ice to Woods Hole Oceanographic Institution for FISH analysis and community determination.

Screen and rubble field sample processing proceeded in much the same manner, except that coral rubble was vigorously scrubbed to remove as much of the epiphytic community as possible. At the March sampling time point, three additional screens were placed for 72 hours at each pool, collected and backwashed into tissue culture flasks, kept at room temperature, and then transported back to WHOI for cell isolation and culture establishment. Filtered seawater for processing was collected on site and passed through a 20µm sieve before use.

Culture Isolation

Unfixed samples were transported to WHOI at room temperature. Cultures were initiated from field samples using previously outlined methods (Guillard & Morton, 2003). Individual cells were mouth pipetted and rinsed three times with filtered seawater, and then placed into individual wells with 1ml of 50% K media, 50% filtered seawater. After a significant increase in cell number occurred, (~15-20 cells total), well contents are transferred to a 25ml tissue culture flask to initiate larger culture in 100% K media.

Culture Species Identification

To assess the taxonomy of *Gambierdiscus* spp., an initial six cultures from sites 1-5 were sequenced to determine their species identity (see following method). Since these cultures consisted of *G. australes* and *G. caribaeus* species, the remaining cultures were preserved and screened using probes detecting *G. australes* and *G. caribaeus*. Cultures that did not label with these two probes were sequenced for species identification. The D8-D10 LSU rRNA region from each *Gambierdiscus* isolate was PCR amplified using primers FD8 and RB (Chinain, Faust, & Pauillac, 1999) and sequenced to determine species identity. Sequences were aligned and consensus sequences were compared with those deposited in GenBank using BLAST sequence similarity searches (National Center for Biotechnology Information).

Temperature Gradient Bar Analysis

Three cultures were used in physiological experiments to test growth rates across different temperature regimes and determine temperature tolerances of two of the most abundant species at Wai'Ōpae in isolated cultures: *G. australes* and *G. caribaeus*. Experiments were performed in triplicate over a temperature range of 16 to 35 °C using two temperature-gradient bars, and growth rates were measured by in vivo fluorescence using a 10-AU Fluorometer (Turner Designs, USA) (Watras, Chisholm, & Anderson, 1982; Xu et al., 2016). Fluorescence measurements (log) over time were plotted to calculate the growth rate from the exponential portion of the curve using the following equation from (Guillard, 1973):

$$\mu \text{ (division day}^{-1}) = (\ln (N_1 / N_0)) / (0.6931(t_1 - t_0))$$

N_1 and N_0 represent fluorescence at times t_1 and t_0 and are used to calculate μ (division day^{-1}), the growth rate. Four sequential transfers of the culture in total took place, where the initial transfer to the temperature gradient bar was used for acclimatization of the culture, and the subsequent three transfers were used for growth rate data generation. Prior to measuring fluorescence, cultures were mixed by hand, or using air blown from a glass pipette tube. Fluorescence was measured twice a week and cultures were transferred when they passed a threshold of 50 relative fluorescence units. Two strains of *G. australes* and one strain of *G. caribaeus* were analyzed. The *G. caribaeus* strain and one of the *G. australes* strains were from the 'Control1' and 'Control2' sites in Wai'Ōpae which are further seaward, and one, a *G. australes* strain, was from the 'Top Lagoon' site which is located closer to shore, and experiences a higher degree of thermal variability.

Species Identification by Fluorescence in situ Hybridization (FISH) Probes

As described in Chapter 3, species-specific probes were developed to detect commonly observed *Gambierdiscus* species in field samples. Probes detecting *G. australes*, *G. belizeanus, G. caribaeus*, *G. carpenteri*, and the *G. polynesiensis* clade were used. Probes were hybridized to one of three fluorophores (Alexa Fluor® 488, 532, or 594) if used in triplicate or one of two fluorophores (Cy3 or Alexa Fluor® 488) if used in duplicate and tested against field samples in conjunction with one or two other probes (Table 4.1). These combinations of two or three probes were used to label fractions of each field sample in a whole-cell hybridization assay, as described in chapter 3 and (Anderson et al., 2005). Calcofluor White was added in a final wash step to enable *Gambierdiscus* identification at the genus level due to its labeling of dinoflagellate thecal plates (20ul Calcofluor White in 15ml 0.2 SET wash solution). Up to three different fluorophores were enumerated in the same sample by flipping between filters on the fluorescent microscope after identification of a *Gambierdiscus* cell through Calcofluor labeling (Table 4.1). At least 2ml of artificial substrate field sample was analyzed from each replicate by each probe. For rubble samples, the sample volume used in each hybridization reaction was adjusted to enable visualization of *Gambierdiscus* cells among other debris. Each hybridization was performed in duplicate at a minimum.

Pacific						
Group	**Probe**	**Target Species**	**Sequence**	**Tm**	**Filter**	**Fluorophore**
1	Gcarib_D1D3	*G. caribaeus*	TGAGACCCACATGTGGAGATTC	**52**	FITC	AF488
	Gaust_D1D3	*G. australes*	TGCCAATCCAGTTGTGTATCTC		Cy3	Cy3
2	Gcarpent_D1D3	*G. carpenteri*	TGATGTAACGCAAGACGCACAG	**53**	FITC	AF488
	Gbeliz_D8D10	*G. belizeanus*	AGATCAGTACGCCAGAGTGACTA		Cy3	AF532
	Gpoly_D8D10	*G. polynesiensis clade*	CTCCGCCAGTGACGTTAAGTAG		TxRd	AF594

Table 4.1 Probe Combinations used in Pacific

Probe name, target species, nucleotide sequence, hybridization temperature (Tm), filter, and fluorophores used in this study.

ANOVA Testing of Artificial Substrate Deployments Time (24 verses 72 hours)

One-way ANOVA testing was use to compare overall *Gambierdiscus* abundance between 24 and 72 hour screen deployments; individual hybridization measurements were grouped within each site as well as grouped within site and month combinations. In every case the null hypothesis of no difference in means between the two groups was unable to be rejected, therefore they were grouped together for remaining statistical analysis.

Supplemental Table 4.1 ANOVA Results by Month

Supplemental Table 4.2 ANOVA Results by Site and Month

Statistical Analysis of *Gambierdiscus* species composition

Species composition data was analyzed using the software package PRIMER6 (Clarke, 1993). For data generated from artificial substrates, species abundance was calculated as cells labeled per ml of labeled field sample. Since probes were not available for all species present in samples (e.g., *G. toxicus* and *G. pacificus*) an 'other' species fraction was also calculated by subtracting the sum of all species abundances from the total abundance of *Gambierdiscus* as a genus, as determined from Calcofluor counts.

For rubble data, species abundance was calculated as cells labeled relative to the total cells in the hybridization reaction. Total abundance of *Gambierdiscus* cells in not quantitative in the case of rubble for this study, since no surface area approximations were attempted; thus, cell abundance quantified on rubble only provide relative community composition. Rubble samples in which no *Gambierdiscus* cells were present were excluded from analysis. Species abundance was averaged over replicates within artificial substrate samples (over 24 and 72 hour screen incubations). Rubble samples for the same pool and sampling time were also averaged.

To assess the similarity of *Gambierdiscus* community composition among sites and over time, averaged abundance data were square root transformed and non-metric Multi-Dimensional Scaling (MDS) plots constructed from a Bray-Curtis similarity index. For these analyses, hierarchical cluster analysis was performed and then mapped onto MDS plots. Each month of sampling was examined individually as well as together as a whole.

Multinomial Modeling of Predicted Relative Species Abundances

In collaboration with Andrew Beet at the Marine Policy Center, the dependence of relative abundance of *Gambierdiscus* species on temperature was examined.

For this model let,

$$[Y_{1j}, Y_{2j}, Y_{3j}, Y_{4j}, Y_{5j}, Y_{6j}] \sim Multinomial(n_j, \underline{p}(x))$$

where the random variable Y_{ij} represents the number of individuals of species i in a sample of size n_j, j=1,…,N such that $n_j = \sum_{i=1}^{6} Y_{ij}$ and $\underline{p}(x)$ is the vector of relative abundances dependent on temperature, x. The multinomial logistic regression model assumes that the log-odds for all species relative to the baseline species be a linear function of the predictor,

$$log\left(\frac{p_i(x)}{p_6(x)}\right) = \beta_{0i} + \beta_{1i}(x) \quad \text{for j = 1, …, 5}$$

where

$$p_i(x) = \begin{cases} \dfrac{exp(\beta_{0i} + \beta_{1i}(x))}{1 + \sum_{k=1}^{5} exp(\beta_{0k} + \beta_{1k}(x))} & for\ i = 1, \dots, 5 \\ \dfrac{1}{1 + \sum_{k=1}^{5} exp(\beta_{0k} + \beta_{1k}(x))} & for\ i = 6 \end{cases}$$

and β_{0i}, β_{1i} for i = 1,…,5 are parameters to be estimated. This model is analogous to the logistic model, except that probability distribution for the response is multinomial instead of binomial. A natural way to address the questions of whether the relative abundances are temperature dependent is to test the null hypothesis, $H_0: \beta_{1i} = 0$ for i = 1,…,5 against the alternative hypothesis $H_1: \beta_{1i} \neq 0$ for at least one species. The likelihood ratio statistic for testing H_0 against H_1 is;

$$LR = 2\left(logL_{H_1} - logL_{H_0}\right)$$

where $logL_{H_1}$ and $logL_{H_0}$ are the maximized values of the log likelihood under H_1 and H_0 respectively. LR is approximately chi-squared distributed, χ^2, with degrees of freedom equal to the difference in the number of parameters under H_1 and H_0. The contribution to the likelihood, $logL_j$ for each sample j is:

$$logL_j\left(\underline{\beta_0}, \underline{\beta_1}\right) = \sum\nolimits_{k \in j} y_{kj} log\left(p_k(x)\right) + \left(\sum\nolimits_{k \notin j} y_{kj}\right) log\left(\sum\nolimits_{k \notin j} p_k(x)\right)$$

where $k \in j$ represents the set of species in sample j for which counts were obtained and $k \notin j$ are the remaining species of which a combined count are available. Note that $\sum_{k \notin j} y_{kj} = n_j - \sum_{k \in j} y_{kj}$.
The full likelihood is the sum of the contributions of all N samples.

$$logL\left(\underline{\beta_0}, \underline{\beta_1}\right) = \sum_{j=1}^{N} logL_j\left(\underline{\beta_0}, \underline{\beta_1}\right)$$

This likelihood can be maximized using standard numerical optimization algorithms.

Results

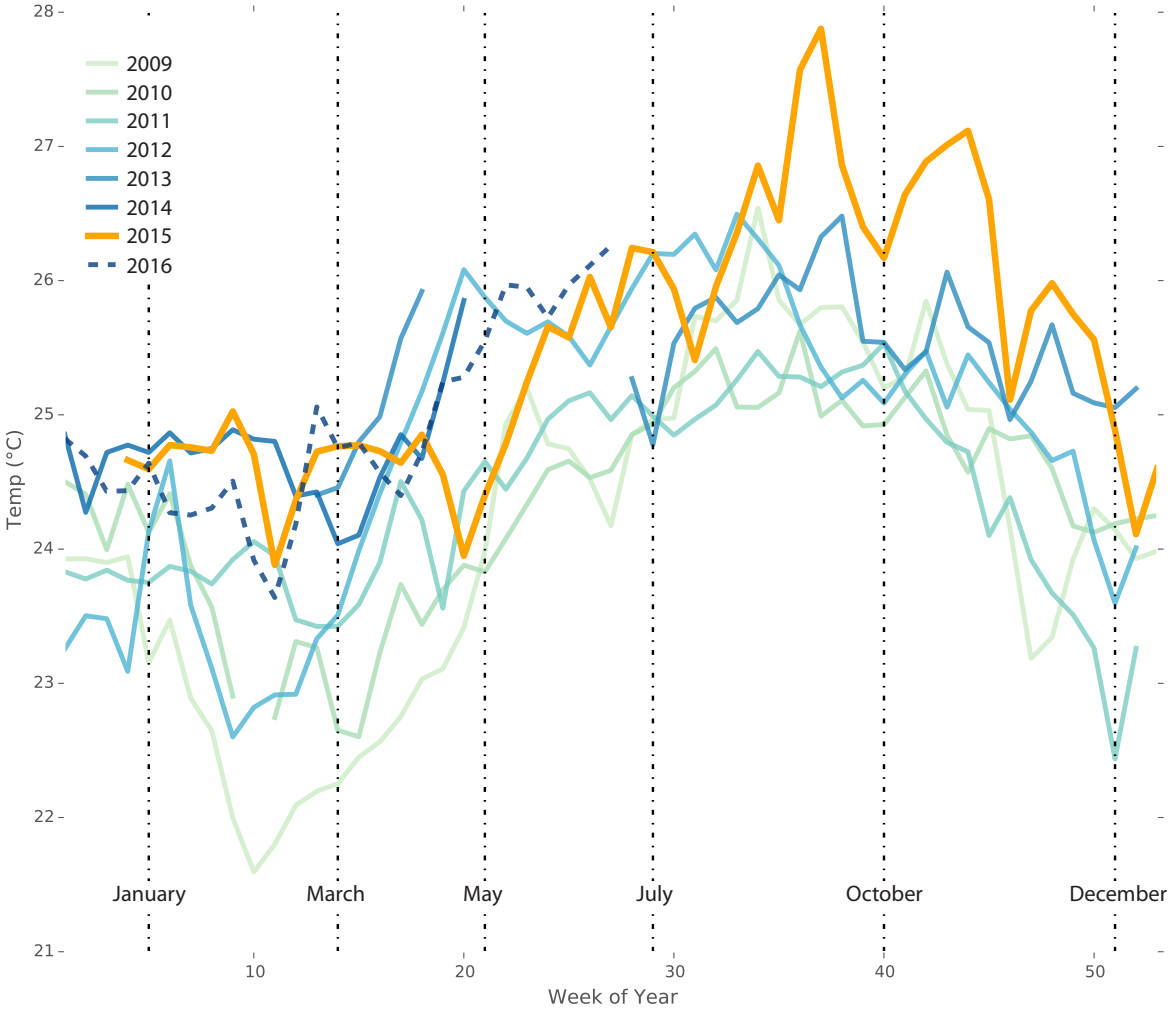


Figure 4.3 SST Values from Hilo, Hawai'i
Weekly average of sea surface temperature (SST) values from Hilo, Hawai'i (NOAA available data); an orange line highlights the year 2015, a blue dashed line highlights data from 2016. Vertical dashed lines represent sampling time points in 2015. The year 2015 had unusually high SST values. Station located at 19° 43.8' N, 155° 3.3' W in Hilo Bay.

Temperature Variability Across Wai'Ōpae

Water temperatures in Wai'Ōpae, Hawai'i, are influenced by influxes of geothermally heated groundwater from the shore and from the seaward side by the ocean. These shifts in temperature are especially apparent at low tides (Figure 4.4) and during higher tides mean temperature differences are minimized (Figure 4.5). From January 2015 to January 2016 pools experienced a range of degrees of temperature variation as well as seasonal shifts in mean temperatures across the site. During three days of sampling in January 2015, the only time point for which a complete

temperature record is available at all sites, site 6 'Hot Pool' and site 1 'Top Lagoon' had the highest temperature variance while site 3 'Control 2' was the most thermally stable (Figure 4.6).

July and October were the hottest months during which sampling took place, with peak mean temperatures occurring in September (Figure 4.7). Over the sampling period pools reached a maximum temperature of 30.7°C and a minimum temperature of 22.8°C (Figure 4.7). Site 6, 'Hot Pool', was the most variable site with the highest variance in temperature throughout much of the year where data are available (Figure 4.9), commonly experiencing daily fluctuations of 2-4°C with a maximum range of almost 5°C. Meanwhile, sites further seaward such as site 5, 'Control 1', more commonly experienced daily temperature fluctuations of 0-2°C, with a maximum range of 3°C. NOAA records of sea surface temperatures from nearby Hilo, Hawai'i, show the pools likely experience a different range of temperatures from nearby ocean waters, especially during summer months (Figure 4.10). In Figure 4.10, site 1 'Top Lagoon' experiences higher temperatures than site 5 'Control 1', which is further seaward.

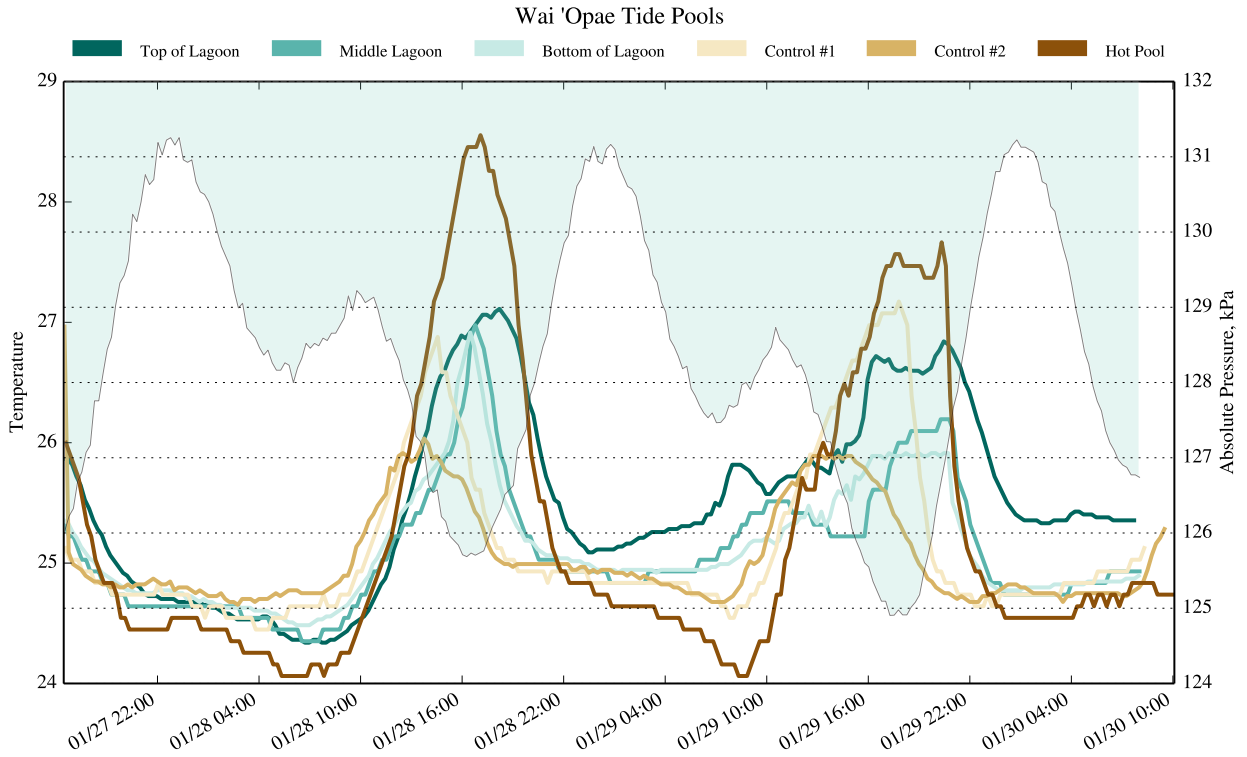


Figure 4.4 January Temperature Variability with Tidal Record
Temperature records from all sites from 1/27/15 to 1/30/15 during sampling in January with corresponding tidal record (pressure record at site 5). With low tide, an incursion of geothermally-heated groundwater raises temperatures at sites in Wai'Ōpae, Hawai'i. Site 6, 'Hot Pool', reaches the highest and lowest temperatures.

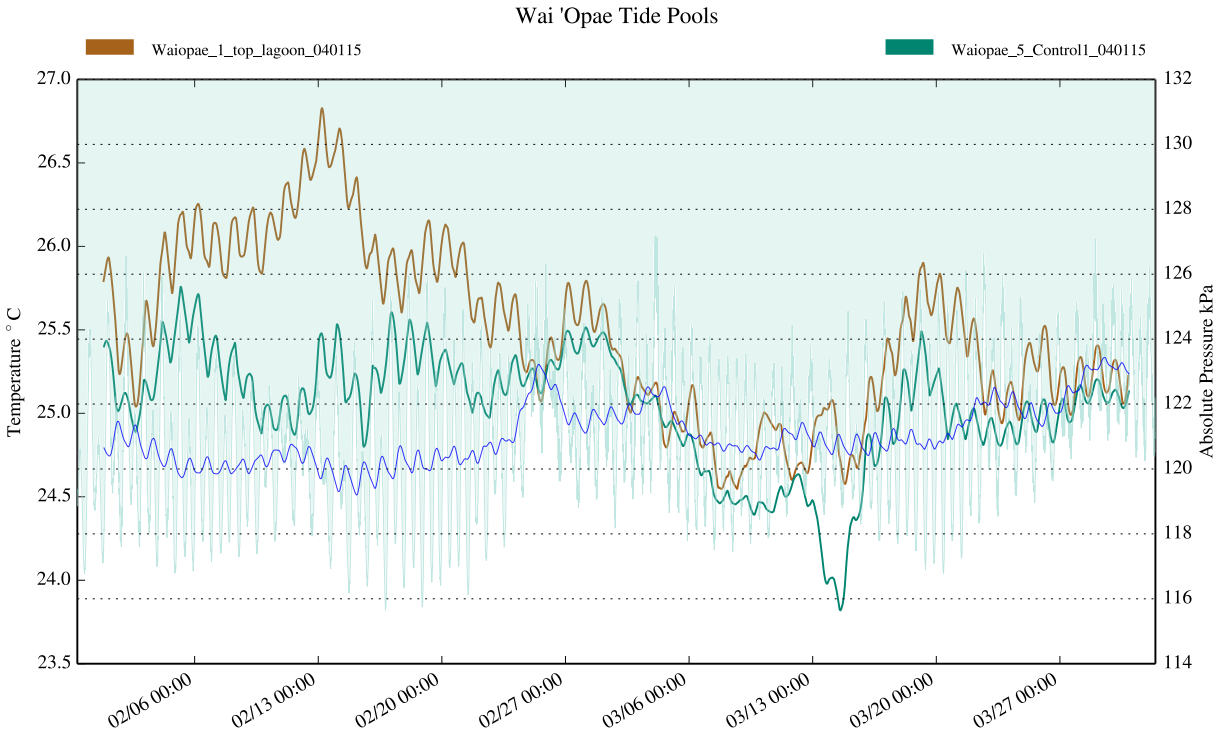


Figure 4.5 Spring Temperatures and Tidal Record
Temperature records from sites 1 and 5, 'Top Lagoon' and 'Control 1', from February to March 2015 with tidal pressure record from site 5. Mean pressure is depicted by the blue line. At mean high pressures (higher tides) more mixing occurs and temperatures at sites 1 and 5 appear more similar. Site 1 is very close to shore while Site 5 is seaward.

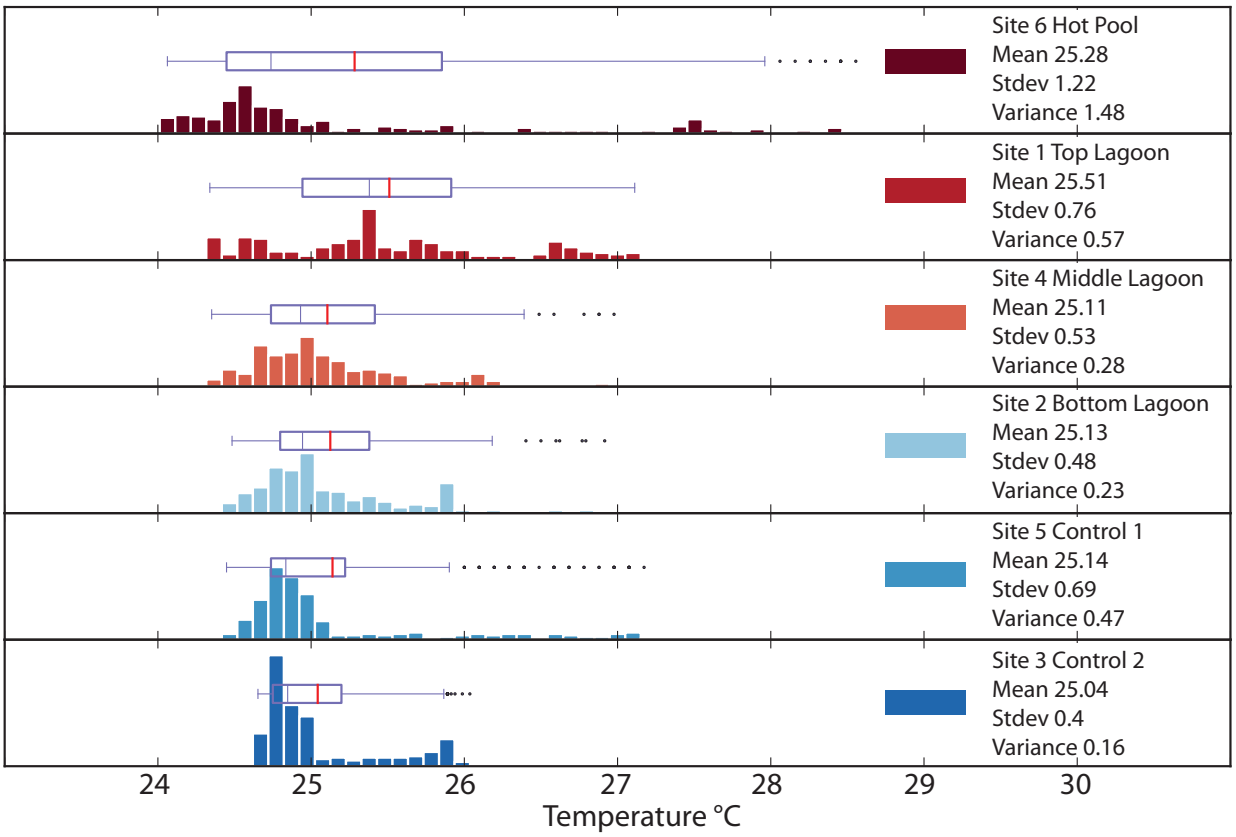


Figure 4.6 Summed January Temperature Variability
Temperature records from every fifteen minutes were binned over the same time period in January (1/27 to 1/30) from each site. Mean and variance of temperature records were calculated (described in rectangle insets and represented by horizontal box plot overlays that show mean temperature (red vertical line), median (thin purple vertical line), a box outlining the interquartile range, whiskers denoting minimum and maximum temperatures, as well as black dots for outlier temperatures). Sites were ordered from closest (site 1 and 6) to furthest from shore (site 3 and 5). Site 6 and Site 1 were the most variable, with the highest variance in temperature.

Salinity Gradient

Salinity (S) ranged from 29.0 at Top Lagoon site to 34.5 at Control 2 at one sampling time point (7/16/15) while at another time point there was no difference (4/1/15, S=35) (Supplemental Table 4.3). Salinity measurements were collected at depth of *Gambierdiscus* substrates (rubble and artificial) with whirl pack bags and measured by refractometer. Previous studies measured salinity in the top fraction of the water column, reporting fresher values in surface waters (Wiegner et al., 2016).

Nutrients in Wai'Ōpae

In a study of nitrogen sources in Wai'Ōpae, total dissolved nitrogen (TDN) along a 100m transect through the lagoon (Sites 1,4, and 2) varied from 18 to 11 μmol L^{-1} with generally higher TDN values closer to shore and lower further seaward (Wiegner et al., 2016) (Supplemental Figure 4.1).

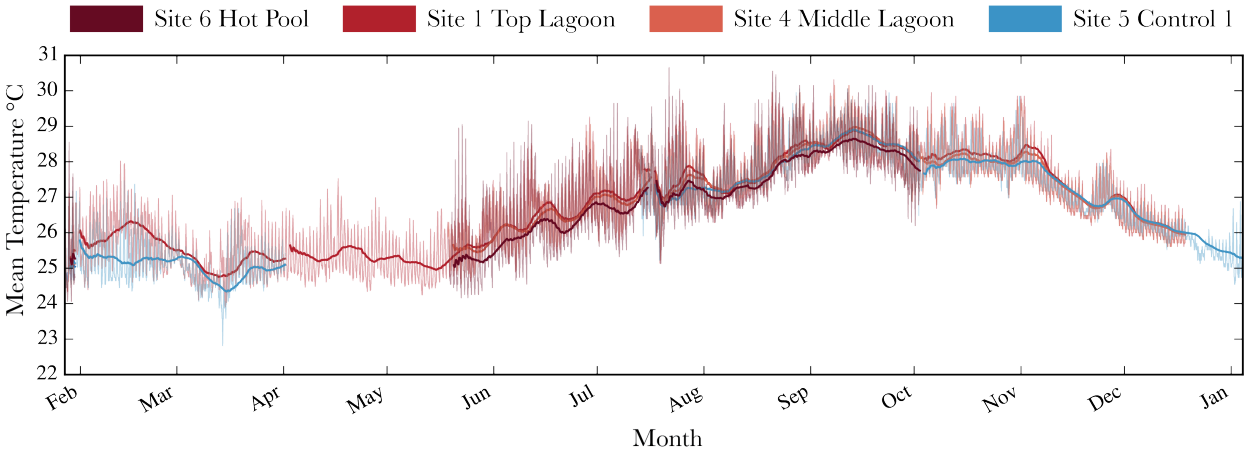


Figure 4.7 Temperature Record Across Pools for Wai'Ōpae
Darker colored lines represent mean temperature of pools. Pools reached a maximum temperature of 30.7°C and a minimum temperature of 22.8°C across sampling dates.

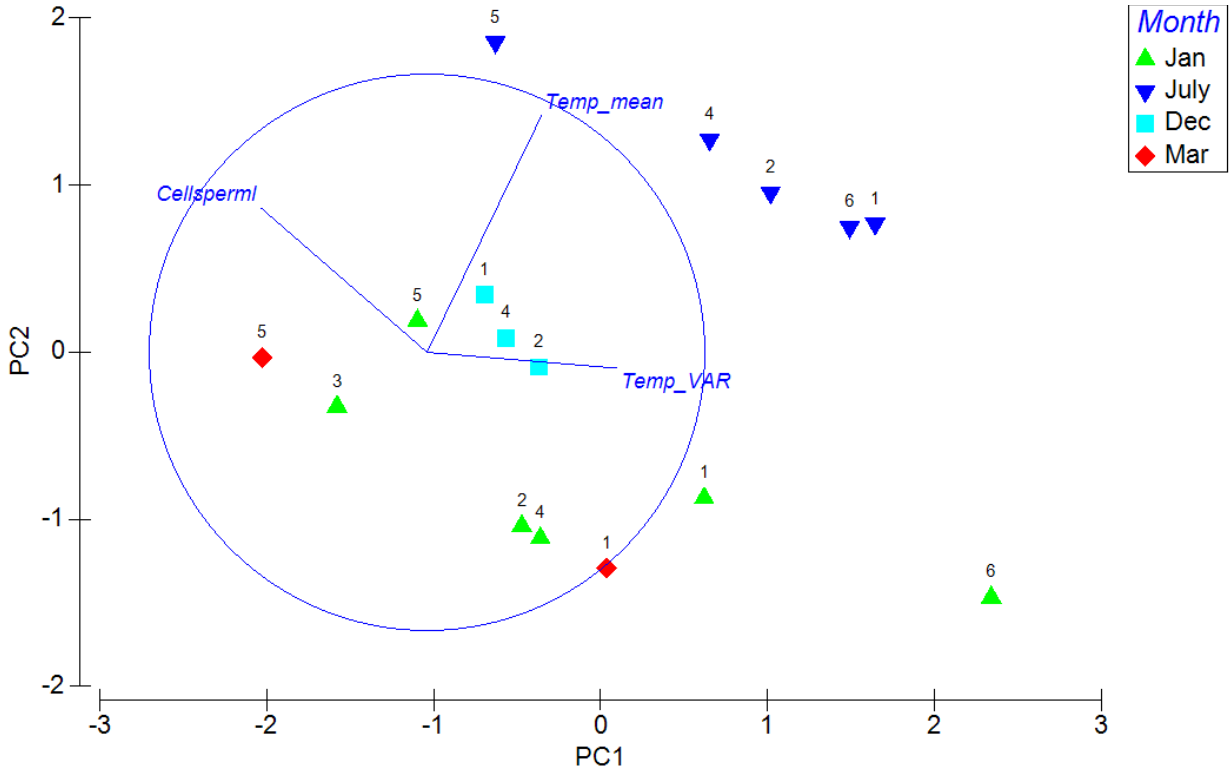


Figure 4.8 PCA Plot of Environmental Factors

PCA plot of mean temperature, variance in temperature, and *Gambierdiscus* abundance in cells per ml sample across sites and months of artificial substrate ampling for which all these data exist. Figure shows the separation of data from site 6, and to a lesser extent site 1, in some months along the 'temperature variance' axis and the separation of data from the month July along the 'temperature mean' axis.

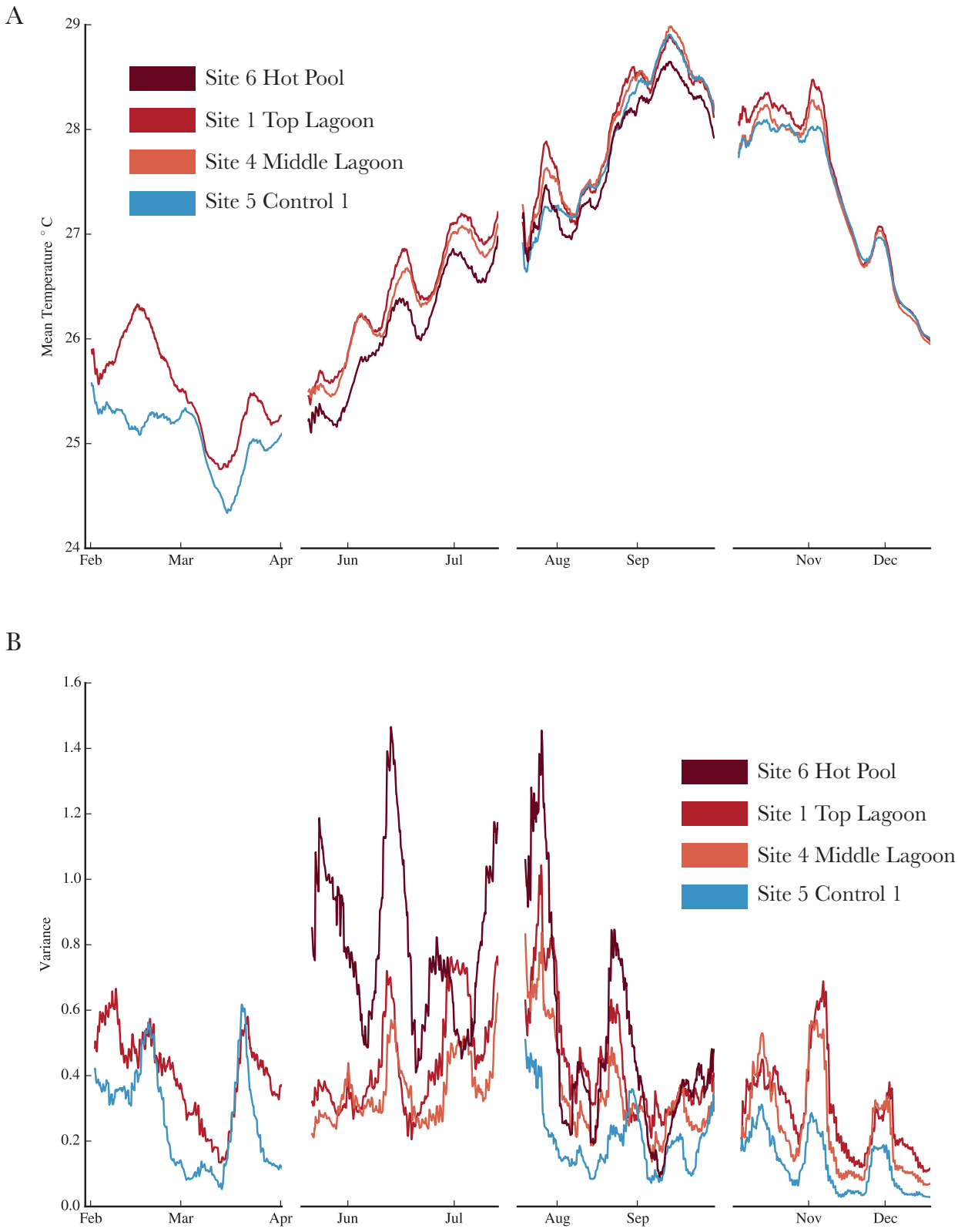


Figure 4.9: Wai'Ōpae Variance and Mean Temperature during Study
A rolling average of temperature records from each site (A) is plotted along with corresponding variance in temperature (B).

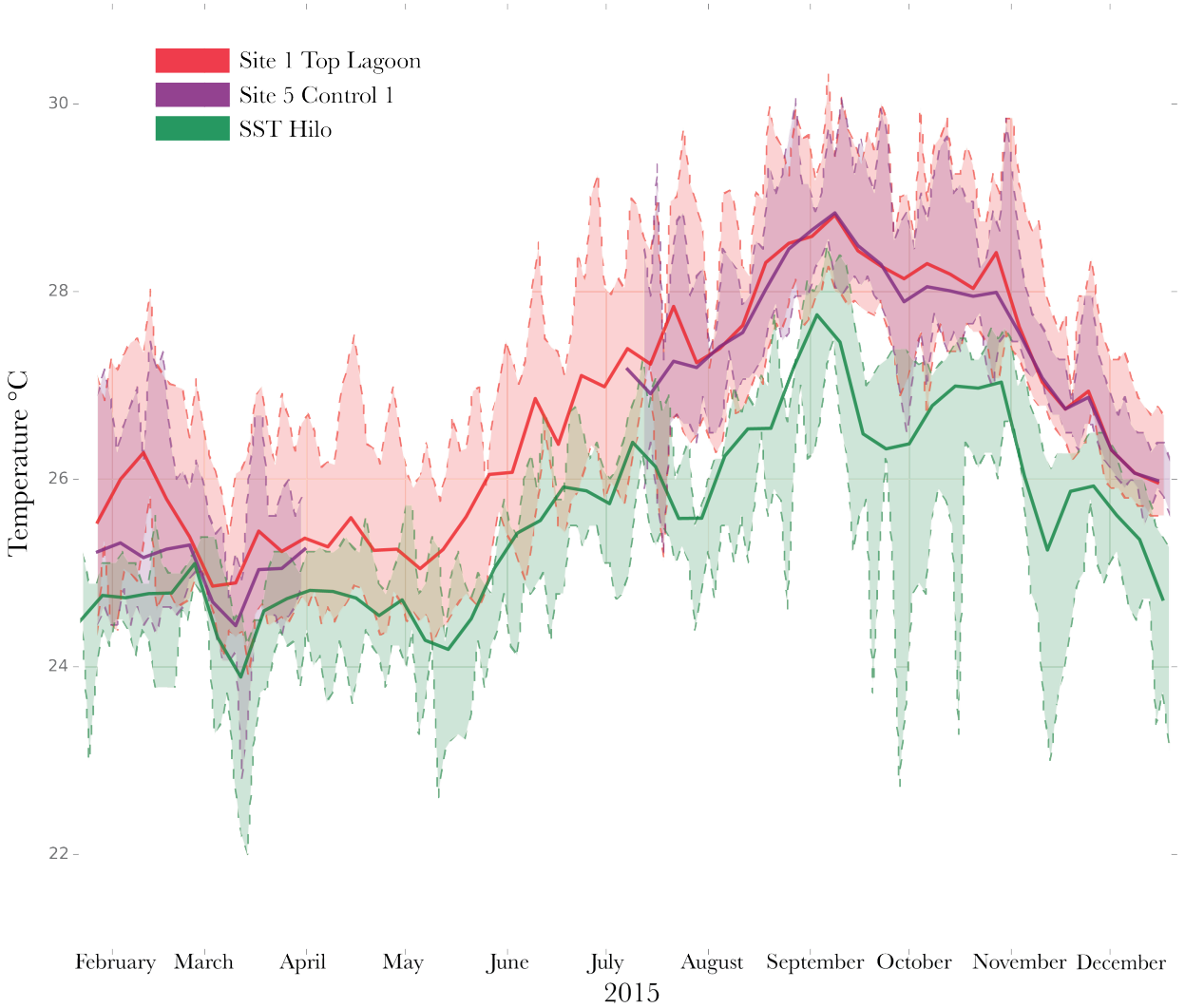
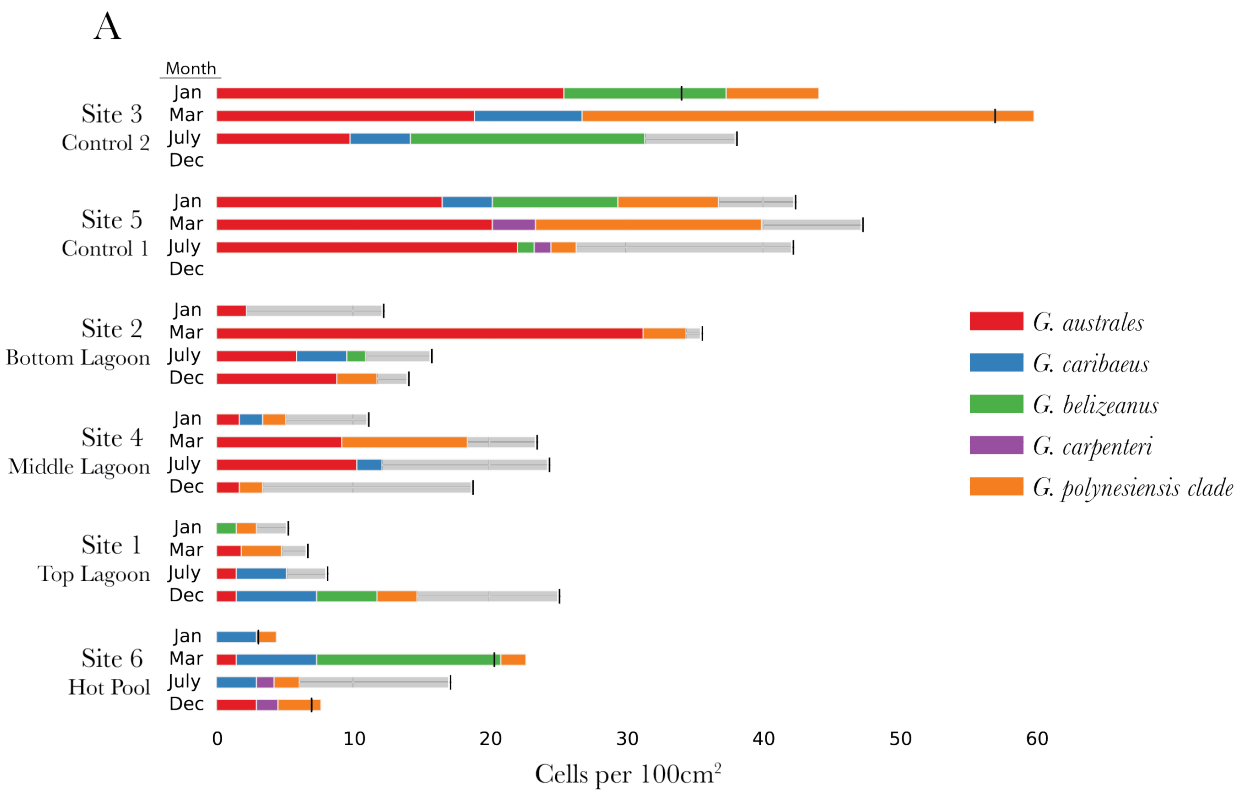
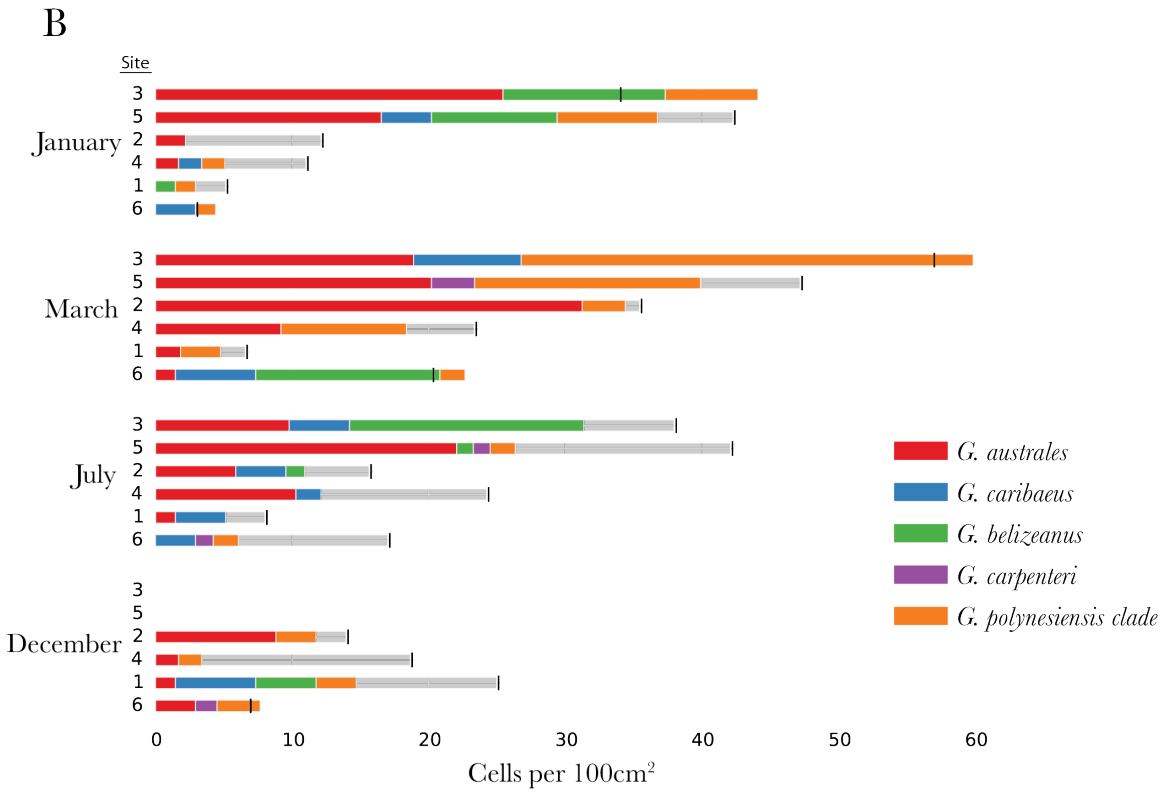


Figure 4.10: **Site 1, Site 5 and Hilo SST Temperature Variability**
Darker colored lines represent mean temperature; shaded area is the difference between the minimum and maximum temperatures experienced (dashed lines). This figure illustrates that pools are more similar to each other than to sea surface temperatures, with site 1 'Top Lagoon' reaching higher temperatures than site 5 'Control 1' which is further seaward.

Figure 4.11 *Gambierdiscus* Species Abundance by Month and Site
Gambierdiscus abundance data from artificial substrates was plotted across the sampling dates. These two joined plots show species abundance grouped primarily by site (A) and by month (B). Black vertical lines indicate averaged *Gambierdiscus* genus abundance over all replicates for the site and month. Colored bars are the averaged *Gambierdiscus* species abundance.


A
Month
Site 3
Control 2
Site 5
Control 1
Site 2
Bottom Lagoon
Site 4
Middle Lagoon
Site 1
Top Lagoon
Site 6
Hot Pool
Jan
Mar
July
Dec
0
10
20
30
40
50
60
Cells per 100cm²
G. australes
G. caribaeus
G. belizeanus
G. carpenteri
G. polynesiensis clade


B
Site
January
March
July
December
3
5
2
4
1
6
0
10
20
30
40
50
60
Cells per 100cm²
G. australes
G. caribaeus
G. belizeanus
G. carpenteri
G. polynesiensis clade

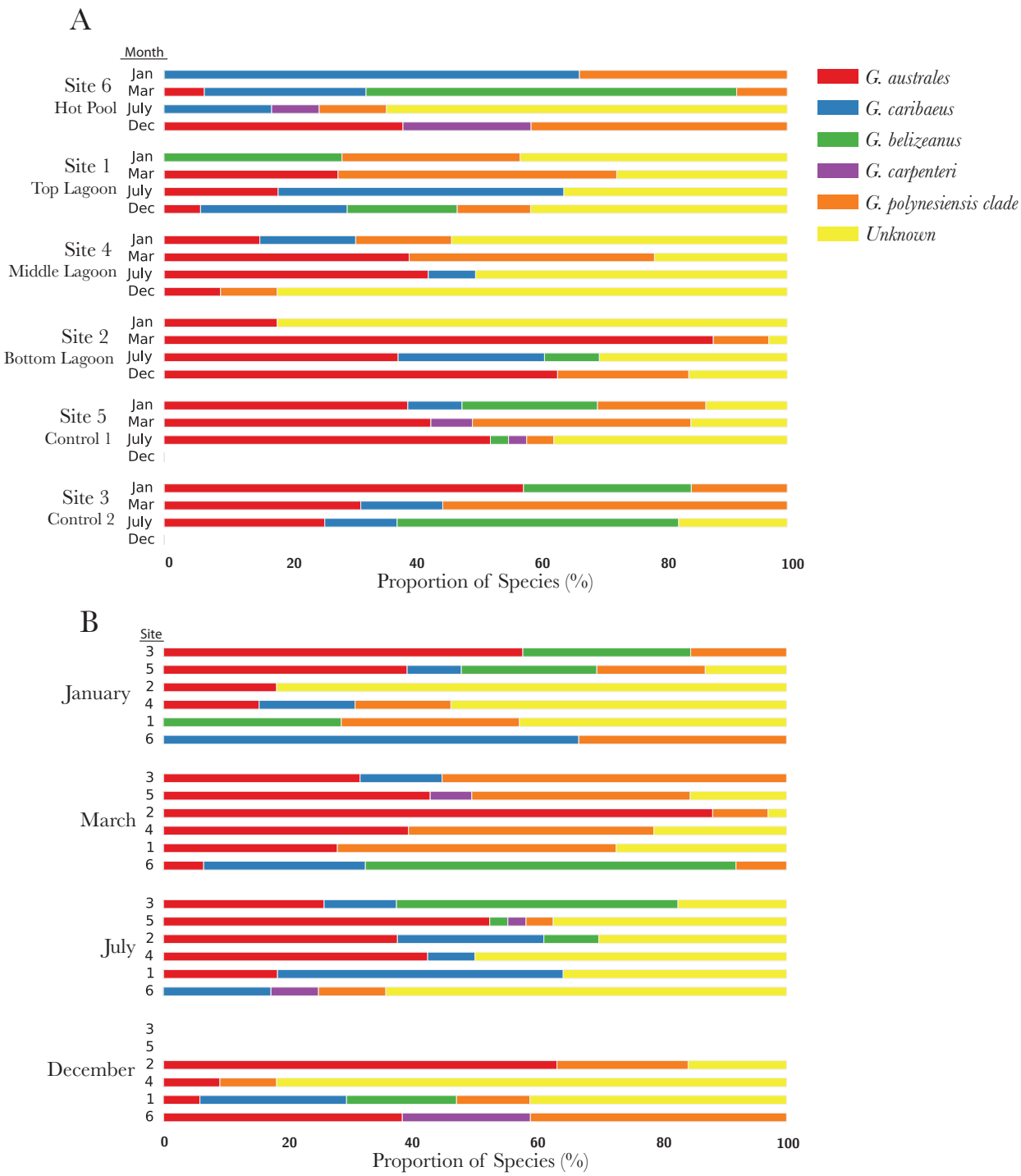


Figure 4.12 *Gambierdiscus* Community Composition by Month and Site from Artificial Substrate Data

Gambierdiscus community composition from artificial substrates was plotted across the sampling dates. These two joined plots show species abundance grouped primarily by site (A) and by month (B).

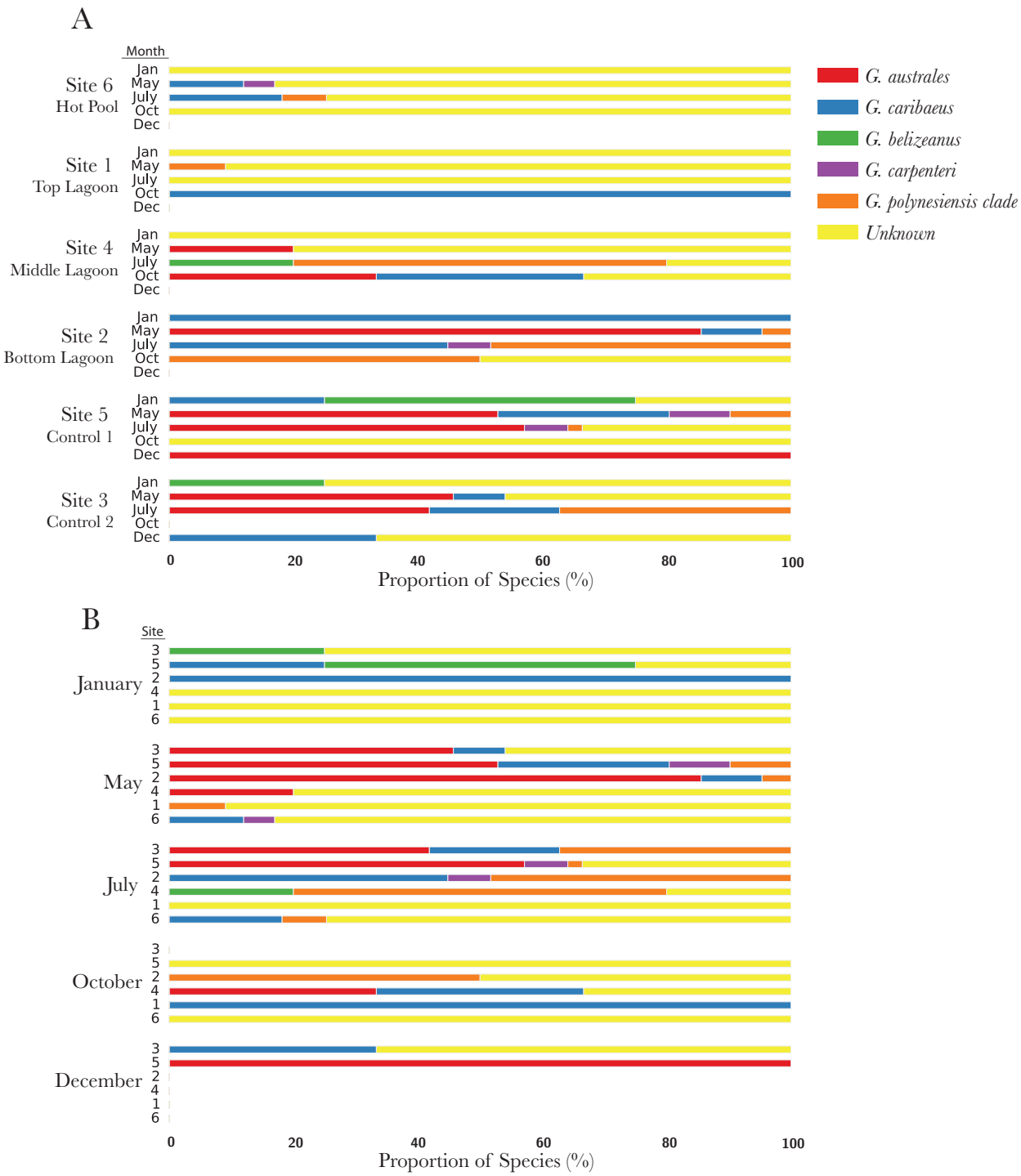
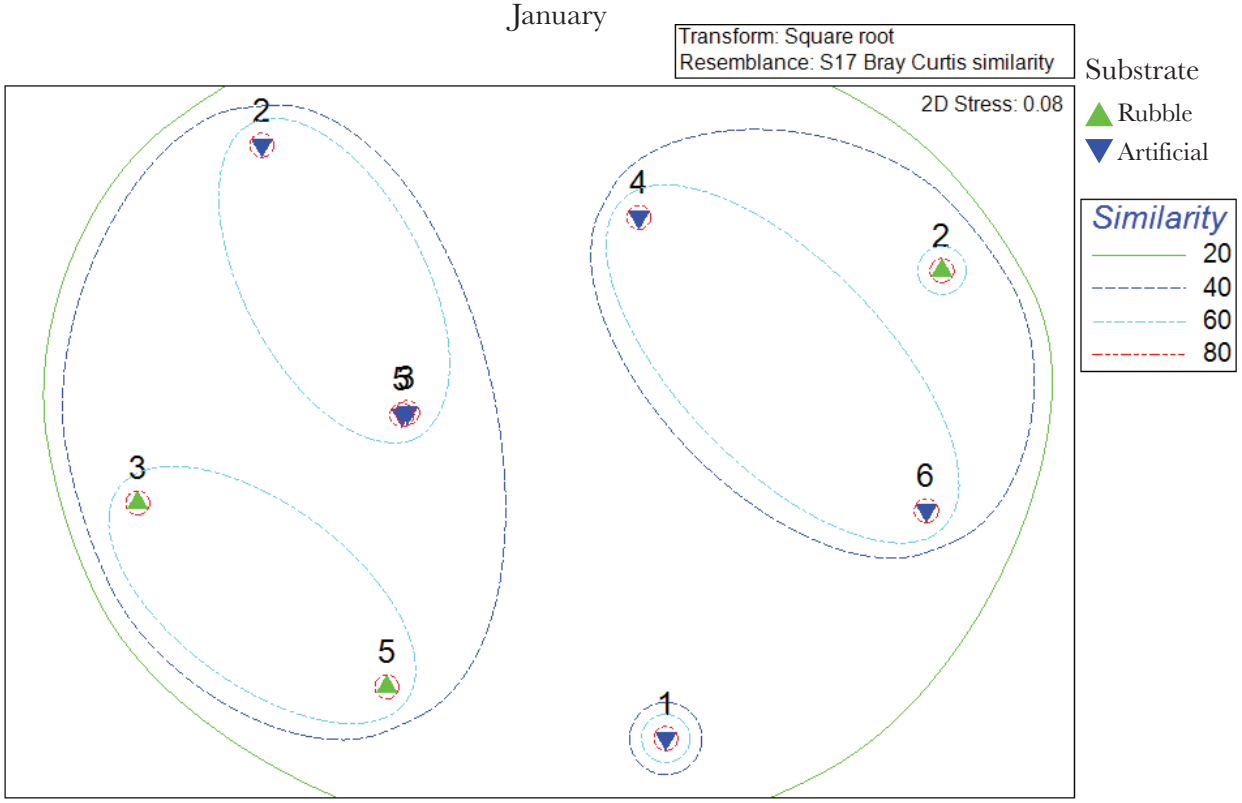


Figure 4.13 *Gambierdiscus* Community Composition by Month and Site from Rubble Data

Gambierdiscus community composition from rubble was plotted across the sampling dates. These two joined plots show species abundance grouped primarily by site (A) and by month (B).

Figure 4.14 Plots of *Gambierdiscus* Community Composition by Month and Site (January and March)
MDS plots of *Gambierdiscus* community composition by month from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid.

January



March

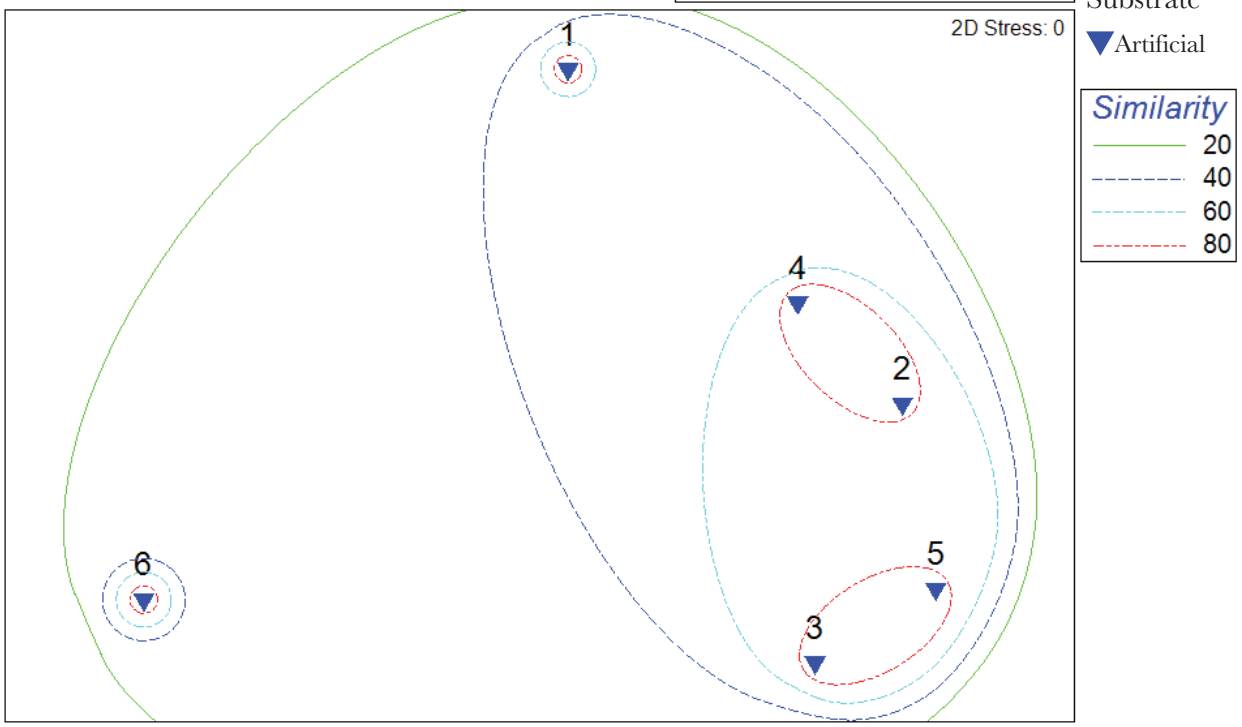
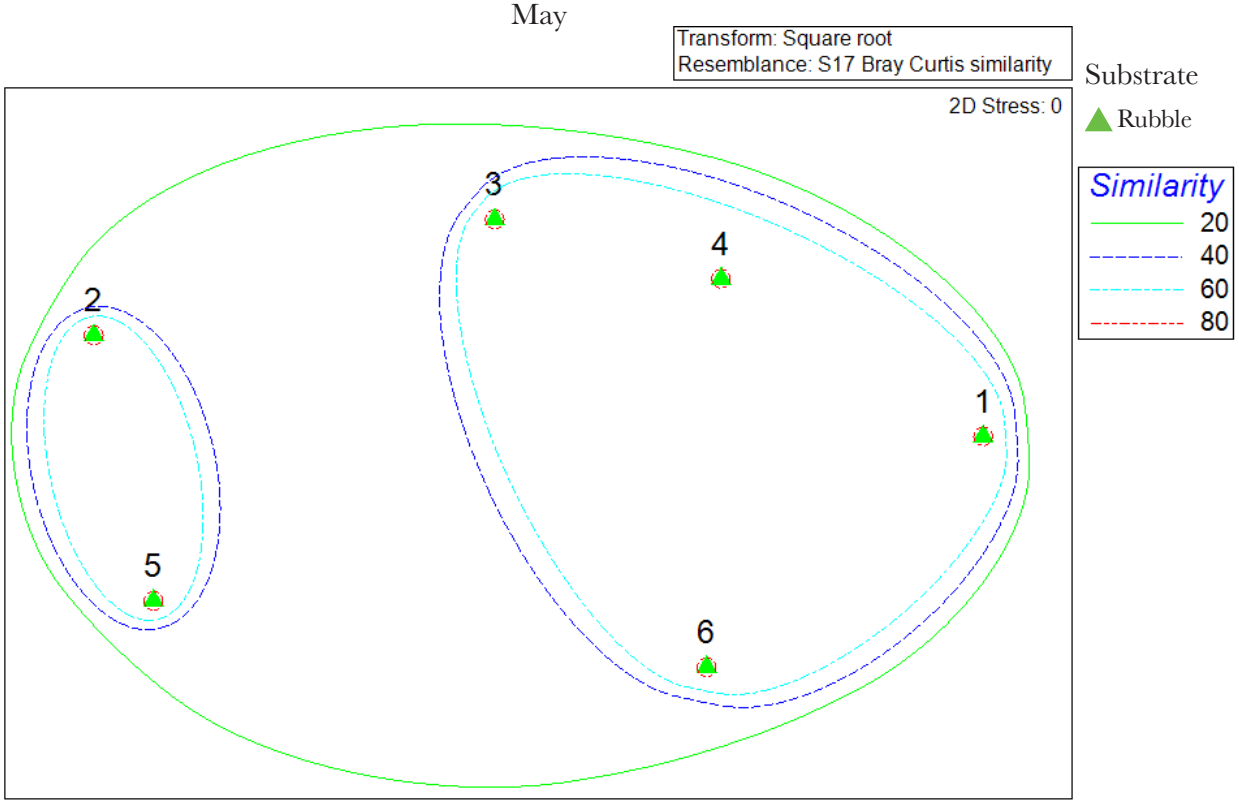


Figure 4.15 Plots of *Gambierdiscus* Community Composition by Month and Site (May and July)
MDS plots of *Gambierdiscus* community composition by month from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid.

May


Transform: Square root
Resemblance: S17 Bray Curtis similarity
2D Stress: 0
Substrate
Rubble
Similarity
20
40
60
80
1
2
3
4
5
6

July

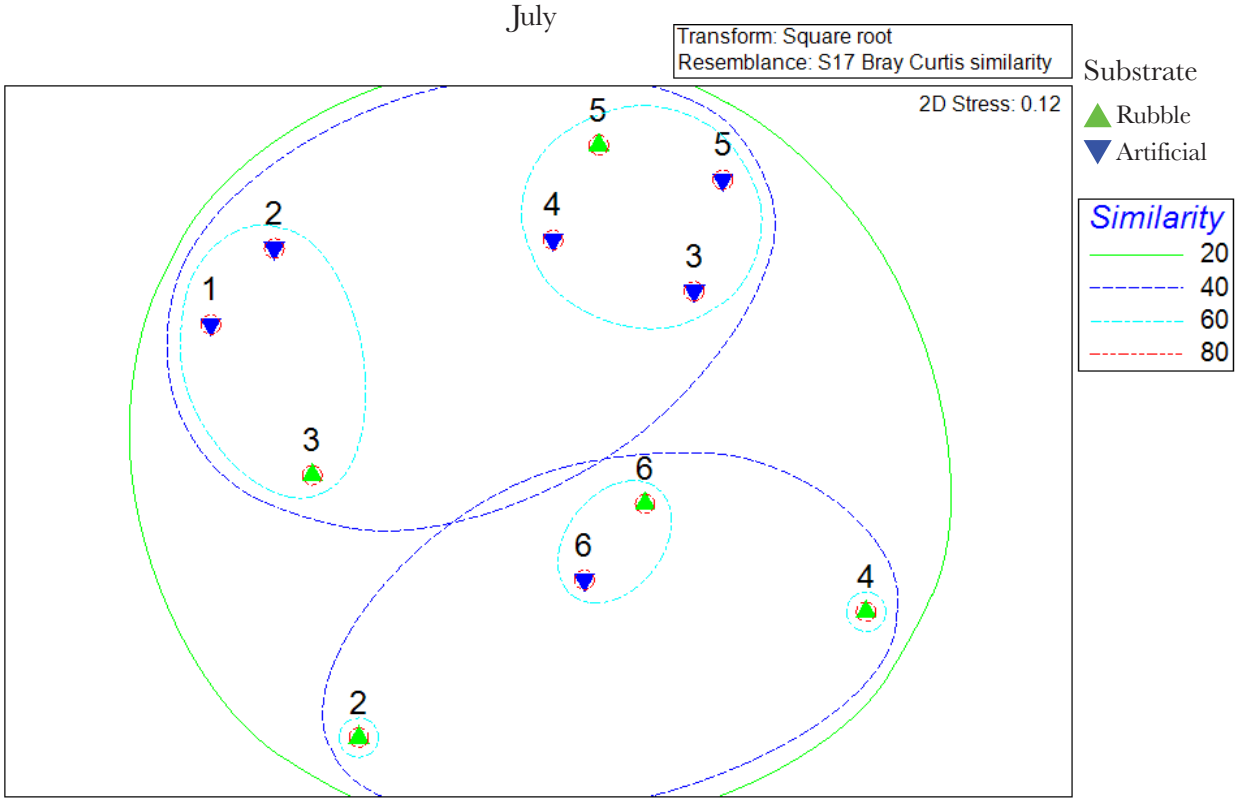
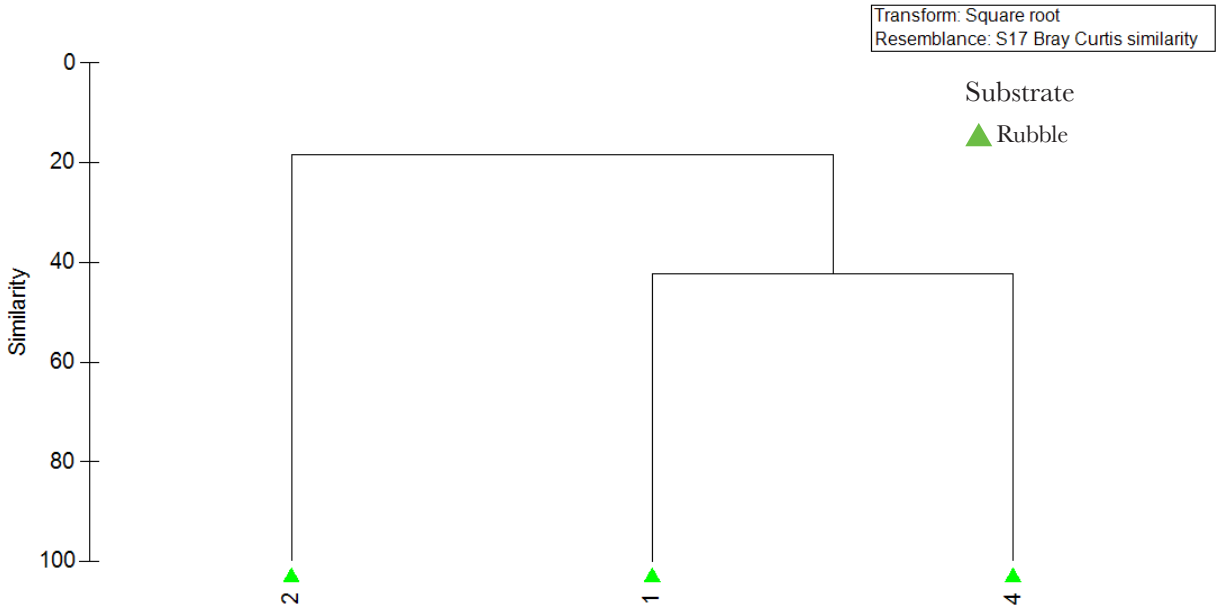
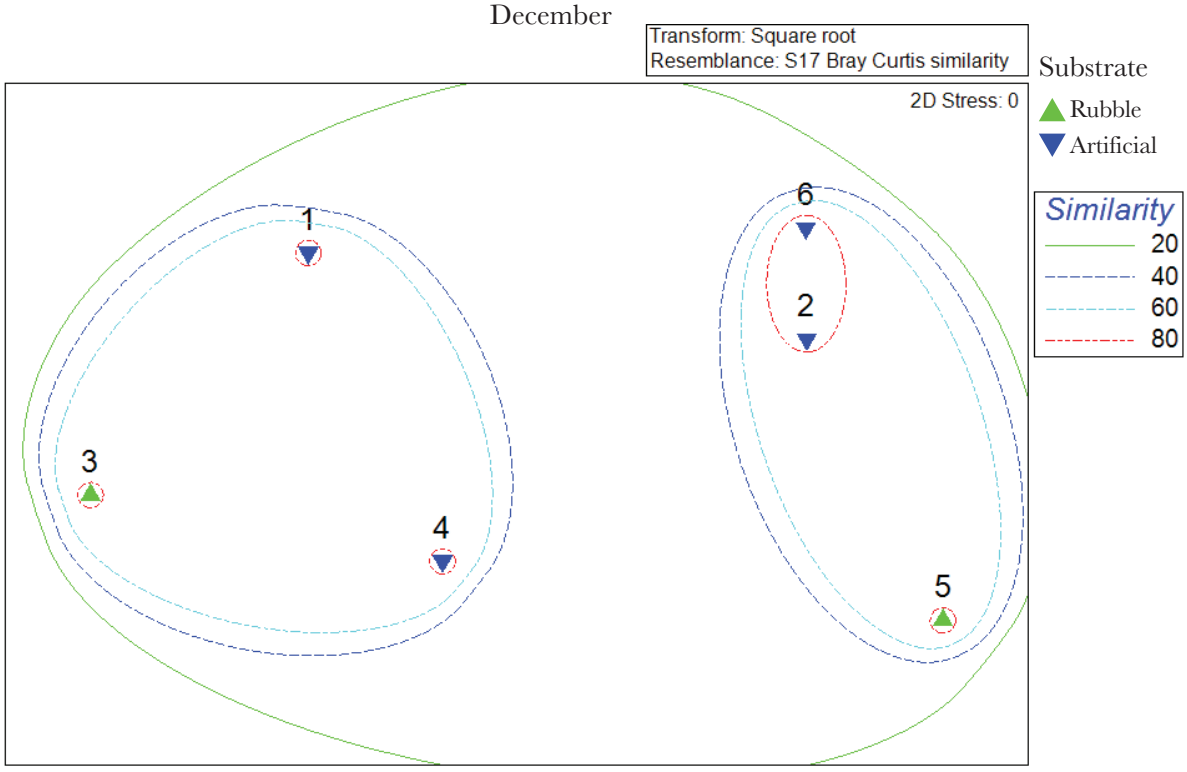

Transform: Square root
Resemblance: S17 Bray Curtis similarity
2D Stress: 0.12
Substrate
Rubble
Artificial
Similarity
20
40
60
80
1
2
3
4
5
6

Figure 4.16 Plots of *Gambierdiscus* Community Composition by Month and Site (October and December)
MDS plots of *Gambierdiscus* community composition by month from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid. In October, too few samples existed where *Gambierdiscus* cells were present to construct an MDS plot. Instead, clustering analysis is shown.

October



December



Patterns of *Gambierdiscus* Genus Abundance

Seasonal shifts of *Gambierdiscus* abundance at the genus level occurred at Wai'Ōpae with the highest cell concentrations occurring in March, followed by July, and then January. Furthermore, there were differences in abundance between sites at the same sampling time point. The two sites located furthest seaward and most exposed to wave action had the highest abundance of *Gambierdiscus* at all screen sampling points (January, March, and July). No screen samples could be taken at these sites in December due to high wave action. The two sites closest to the shore, Top Lagoon (1) and Hot Pool (6), had the lowest *Gambierdiscus* abundances in January, March and July. However in December, of the four sites sampled (Hot Pool, Top Lagoon, Middle Lagoon, Bottom Lagoon), Top Lagoon had the highest *Gambierdiscus* abundance. Artificial substrates showed no statistical difference in *Gambierdiscus* abundance between 72 and 24-hour deployments (ANOVA testing, Supplemental Tables 4.1,4.2).

Patterns of *Gambierdiscus* Species Abundance

There were differences in *Gambierdiscus* community composition seasonally across Wai'Ōpae at a single sampling site, and between sites at the same time point (Figure 4.11; Figure 4.12; Figure 4.13). From artificial substrate data, *Gambierdiscus australes* was a major community member across many of the sites and time points, as well as the *G. polynesiensis* clade. However, while *G. australes* was the most common species at the high abundance sites in March and July, at those times it was much less common at sites 1 and 6 ('Hot Pool' and 'Top Lagoon') which experience a higher degree of temperature variability. Conversely, rubble measurements overall saw less *G. australes* and more 'Unknown' cells. Rubble and artificial substrate sampling methods largely produced different *Gambierdiscus* community composition data (Figure 4.12; Figure 4.13). In data from July, a month in which data from both artificial substrates and rubble is available, sites 5 and 6 had a similar community composition in both artificial substrate and rubble measurements but the remaining four sites had divergent communities based on the two substrate methods.

PRIMER6 analysis

Site 6, or 'Hot Pool', in PCA plots of temperature data appears different from other sites within the environmental data, skewed along the temperature variation axis (Figure 4.8). The month of July is separated from January, March, and December along the mean temperature axis (Figure 4.8). In March, site 6's species diversity is different from the rest of the sites, with Site 1 'Top Lagoon' also sharing less similarity (Figure 4.14). In other months, such as January, the furthest seaward sites, sites 3 and 5, can be very similar in rubble and artificial substrate data. Site 6 also displayed seasonal patterns to species diversity with 60% similarity between May and July, 40% similarity between January and March, and December being the most different (Supplemental Figure 4.8). Across all sites except site 1 and site 6, artificial substrate samples shared at least 40% similarity across all months sampled (Supplemental Figures 4.6-4.8). Within samples from site 5 and site 3, artificial substrate samples across months were even more similar (60% similarity) (Supplemental Figures 4.7 and 4.8).

Predicted abundances of *Gambierdiscus* species from Multinomial Model

Observations were grouped by temperature (irrespective of site or month designation) and species abundances were modeled over a range of five different temperatures experienced at Wai'Ōpae.

As modeled temperatures increased from 25 to 27 degrees, relative abundance of the *G. polynesiensis* clade, *G. belizeanus*, and *G. carpenteri* decreased while that of *G. caribaeus* and species represented in the unknown fraction increased. *G. australes* has higher relative abundance at 26 and 27 degrees than at 25 and 28 degrees, but modeled abundance remained high throughout the temperature range.

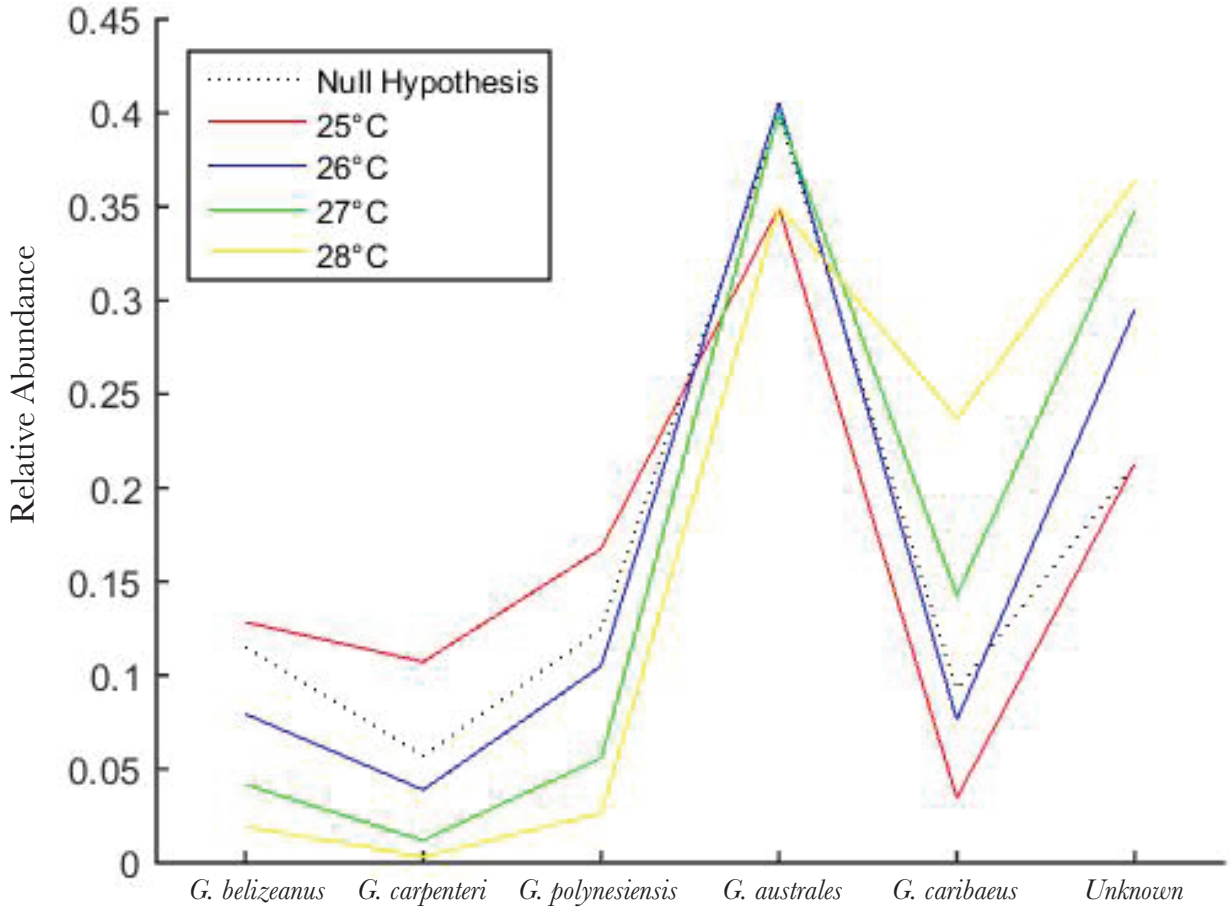


Figure 4.17 Relative Abundance of *Gambierdiscus* Species
Fitted relative abundances of *Gambierdiscus* species under $H_1: \beta_{1i} \neq 0$ as a function of temperature.

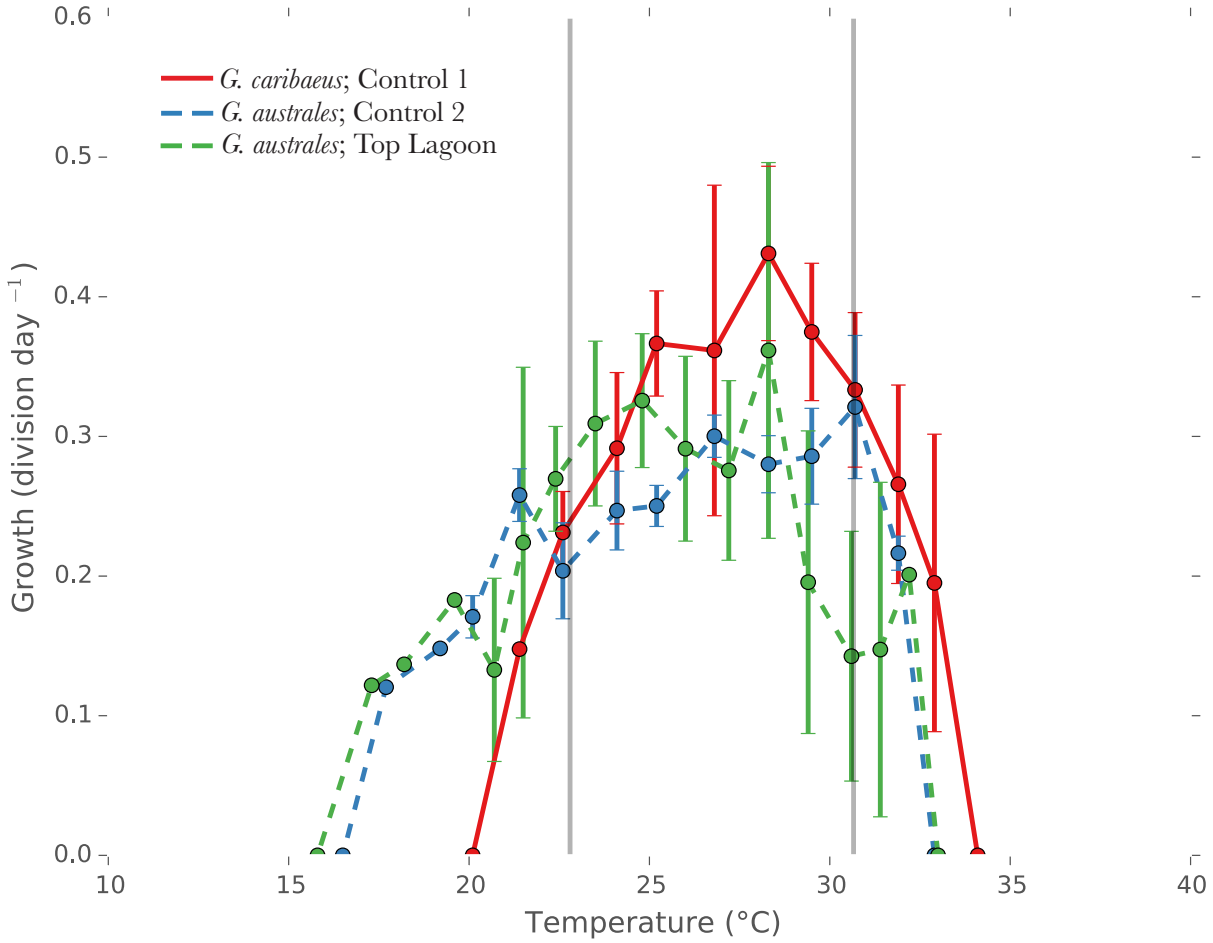


Figure 4.18 Temperature and Growth of *G. caribaeus* and *G. australes*

Growth rate as a function of temperature for one *G. caribaeus* and two *G. australes* strains from Wai'Ōpae. Vertical grey lines indicate lowest and highest recorded temperature experienced at Wai'Ōpae in 2015. Data points without error bars represent temperatures at which only one replicate was able to grow to sufficient density to calculate a growth rate. Therefore, they represent possible growth of a strain but neither favorable nor consistent growth. These growth rates are under optimum nutrient and light conditions.

Temperature Gradient Bar Analysis

Gambierdiscus caribaeus and *G. australes* strains from Wai'Ōpae had similar growth ranges with *G. australes* able to grow at slightly lower and *G. caribaeus* able to grow at slightly higher temperatures (Figure 4.18). *Gambierdiscus caribaeus* has a higher growth rate within the range of temperatures seen at Wai'Ōpae. There is some strain variation between the two *G. australes* strains isolated from different environments: *G. australes* from site 1 'Top Lagoon' had a lower growth rate at high temperatures and a higher growth rate at lower temperatures seen at Wai'Ōpae compared to *G. australes* from site 5 'Control 1'.

Gambierdiscus Culture Identification

In total 56 cultures were isolated and characterized: 52 isolated from samples taken from artificial substrates deployed in March 2015 and vie isolated from rubble samples collected in October 2015 (Table 4.2). Cultures were mostly identified as *G. australes* and *G. caribaeus* (37 and 15 cultures respectively) with two cultures of *Gambierdiscus* species type 4 and one each of *G. silvae* and *G. scabrosus*. To my knowledge this is the first report of *G. silvae* isolated from the Pacific Ocean. *Gambierdiscus* species type 4 and *G. silvae* are both in the *G. polynesiensis* clade and are detected by that probe. *Gambierdiscus scabrosus* was not detected by FISH probe in this study and is represented in the 'Unknown' fraction of cells.

					***G. polynesiensis* clade**		
Site	**Month**	**Total**	***G. australes***	***G. caribaeus***	***G. sp. type 4***	***G. silvae***	***G. scabrosus***
Top Lagoon	March	9	6	0	2	0	1
Middle Lagoon	March	12	8	4	0	0	0
Bottom Lagoon	March	9	7	1	0	1	0
Control 1	March	8	8	0	0	0	0
Control 2	March	13	4	9	0	0	0
Control 1	October	5	4	1	0	0	0
	Total:	56	37	15	2	1	1

Table 4.2 Species Identity and Isolation Location of Cultures

Cultures were isolated from field samples taken in March and October of 2015. Table lists total number of clonal cultures generated from single-cell isolations, their species designations, and collection site.

Discussion

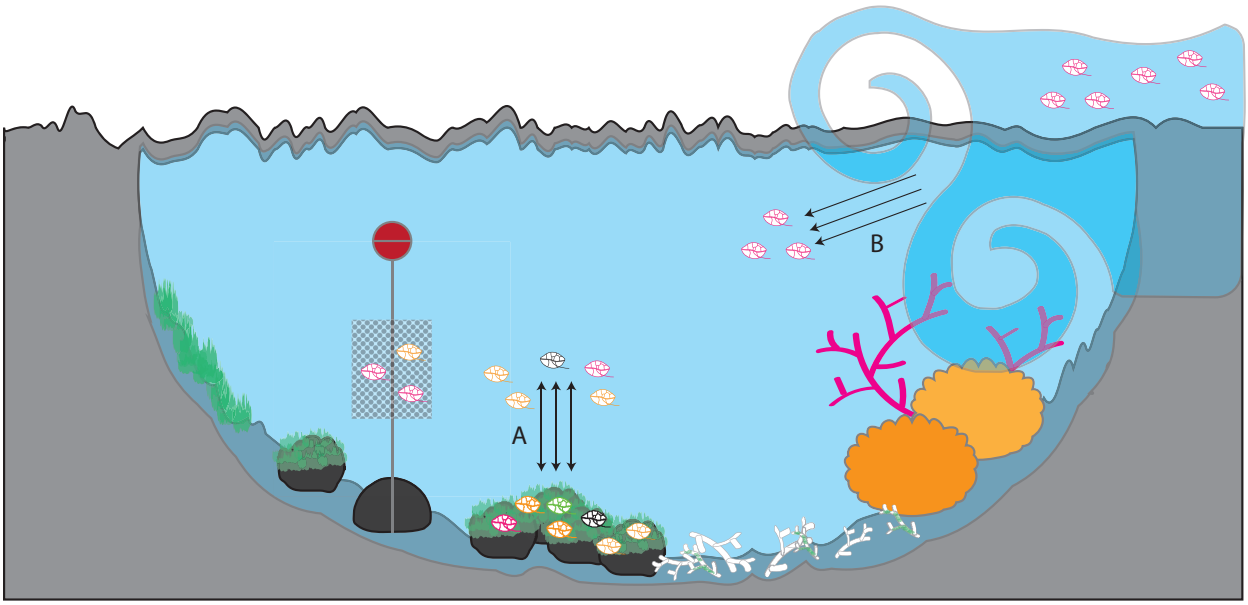


Figure 4.19 Conceptual Figure of Wai'Ōpae Site
Artificial substrates (screens) measure *Gambierdiscus* in the water column. These cells may be in equilibrium with cells on natural substrates such as turf algae (A) or may be influenced by cells immigrating from other environments through water circulation (B). Colors represent different species of *Gambierdiscus*. Cells that are growing on natural substrates are subject to local selection pressures while immigrating cells reflect pressures in their source environment. In the case of Wai'Ōpae, pools locally experience different selection pressures (including temperature) than the surrounding environment. Some species may be favored under different conditions, leading to their increase in natural substrates and subsequent increased representation on artificial substrates.

A guiding motivation of this study was to determine the effect of changing environmental variables on *Gambierdiscus* abundance and diversity. At Wai'Ōpae, a series of tide pools experience different ranges of temperature (differences in temperature variability) on a daily basis while mean temperatures across the site shift seasonally. Using the rRNA-based probes developed in Chapter 3, for the first time it was possible to document seasonal and intersite differences in *Gambierdiscus* abundance and diversity through FISH labeling of artificial substrate and rubble samples, and isolation of cultures from field samples. Using field abundance and temperature measurements, the relative species abundance at Wai'Ōpae was then modeled under different temperature regimes (Figure 4.17), showing that *G. australes* has the highest predicted abundance over a range of mean temperatures, *G. caribaeus* has highest abundance at the highest temperatures, and the *G. polynesiensis* clade, *G. belizeanus*, and *G. carpenteri* have highest abundances at lower mean temperatures. Recent papers have predicted an expansion of *Gambierdiscus* as sea surface temperatures rise above 29°C (P. A. Tester et al., 2010) and a shift in *Gambierdiscus* community structure in the Caribbean to favor *G. caribaeus*, *G. belizeanus*, and *Fukuyoa ruetzleri* with a decline in *G. carolineanus* and *G. ribotype 2* under future climate change scenarios (Kibler, Tester, Kunkel, Moore, & Litaker, 2015). These studies are based on historical rates of

CFP, laboratory-measured growth rates for *Gambierdiscus* under a range of temperatures, and predicted future temperatures for the Caribbean and Gulf of Mexico. Here, the conclusions from this study are drawn from field observations of shifting community composition due to thermal variance. Some of these conclusions are complementary to previous studies, such as the modeled higher abundance of *G. caribaeus* at high temperatures, but some are not and suggest that further investigation is warranted.

The high diversity and flux of *Gambierdiscus* species through time at Wai'Ōpae was surprising. Species have been shown to co-occur previously (Nishimura et al., 2013; P. A. Tester et al., 2014) but this is the first published study quantifying *Gambierdiscus* community abundance and composition from field samples in such close spatial proximity through seasons and from multiple substrate types. Novel application of FISH probes to field samples and modeling of results through a multinomial analysis have revealed trends of spatial and seasonal species variation that are informative to determining future risk of CFP.

Low Cell Number Sampling and Analysis Difficulties

Wai'Ōpae was chosen because of its unique range of temperature variability across a semi-connected site and its history of studies of coral reef ecology including *Gambierdiscus*. However, Wai'Ōpae also lies on the windward side of the island of Hawai'i that has lower *Gambierdiscus* abundance in general, so cell counts were low throughout this study. It is a common observation that *Gambierdiscus* populations are higher, and the risk of toxicity in fish greater, on the leeward side of islands.

Another constraint to this study was that it was only possible to use three FISH probes in one hybridization reaction because of filter limitations using fluorescent microscopy; only three species could thus be enumerated in the same hybridization reaction. Accordingly, at low cell numbers, it becomes less likely that a probe will detect target cells, and thus that there will be "unknown" cells that were not labeled by the three probes used in a single reaction. Through multinomial modeling of relative abundances of species, observations were used to predict whether the unlabeled 'Unknown' fraction observed from one probe set was likely to consist of species that would have been detected with the alternative probe set, or if those cells represent a species for which no probe is yet available. Another possibility is that unlabeled cells may be compromised and unable to be labeled by FISH probes due to low quality or low abundance of rRNA. Although only *Gambierdiscus* cells that displayed autofluorence were enumerated (empty thecal plates were not included in cell counts), future analysis using a universal positive probe in combination with species-specific FISH probes would allow discrimination of this effect.

Substrate Sampling Bias

The standardization of abundance measurements has long been known to be a problem in traditional sampling of *Gambierdiscus* and other epiphytes through macroalgal harvesting. Typically cells are enumerated in terms of grams wet-weight of macroalgae, a measurement that varies widely when comparing different species of macroalgae that have different surface area to weight ratios. A better measure would be to take into consideration the surface area available to *Gambierdiscus* for colonization, however these measurements are difficult to make (Lobel et al., 1988).

Different sampling strategies have varying biases towards *Gambierdiscus* abundance estimates at a given site. Traditional rubble and macroalgal sampling (where the entire substrate or a subsection is collected, shaken or scrubbed, and then filtered) represents the best approximation of *Gambierdiscus* species composition on the reef when considering risk from herbivorous fish, and is relatively consistent in terms of abundance over the same macroalgal species. However, this method is constrained by difficulties in obtaining these substrates (some sites may be limited in algal cover perennially or during certain times of the year) and presents difficulties in comparing abundance results across locations and time points (due to difficult surface area approximations of macroalgae and changing substrate availability at a site). Furthermore, there are sample processing difficulties, as *Gambierdiscus* is often a minor community member in a larger benthic assemblage and commonly a high level of detritus is present in samples, making enumeration more difficult.

The strengths of artificial substrate sampling (P. A. Tester et al., 2014) lie in its ease of application to a variety of sites and its putative comparability across sites (due to its fixed and measureable surface area). However, it is still unclear if different flow environments affect the ability of *Gambierdiscus* to settle on this new substrate, and if once settled, *Gambierdiscus* cells grow in abundance as on a natural substrate or will try to emigrate to an algal substrate.

Furthermore there are the biological questions of whether the cells collected on the screen are: a) representative of the local site's diversity and abundance or b) cells that are being transported from one environment to another by currents, and represent the diversity and abundance of *Gambierdiscus* across a broader area (Figure 4.19). If these cells are derived from the surrounding environment, we would also have to assume no species-specific or environment-specific differences in cell behavior: cells would need to move up and down the water column freely in all environmental conditions, and settle on screens (or be trapped there) in the same way they would settle on their natural substrates, with no species-specific differences in this behavior. Although it is likely these conditions are not met, and there are influences from species-specific behaviors as well as immigrating cells, how large of an effect this has on artificial substrate measurements remains to be determined, as well as whether or not these measurements still represent local diversity overall.

These questions are particularly important for this study as I am comparing linked sites over a relatively small geographic area, but they are also important in a broader context as *Gambierdiscus* is known to have a patchy abundance on reefs, and as seen in this study, different species can be favored among small microenvironments. Since fish may graze over a broad area, or be relatively localized, different substrate sampling strategies may be more or less appropriate depending on the monitoring goals. Some coral reef fish have well-defined, local grazing areas. Fishing strategies that focus on taking these 'homebodies' would require information on local abundance of *Gambierdiscus* to assess risk. Data on the ability of *Gambierdiscus* to be transferred through the water column would enable a better understanding of this immigration effect on artificial measurements.

Furthermore, it is not yet clear how long artificial substrates should remain in the environment to represent local diversity and abundance. Twenty-four hours has been proposed as sufficient for sampling local abundance (P. A. Tester et al., 2014) and was been supported in this study since

abundance was not statistically different between 24 and 72 hour deployments, but this period of time may not be appropriate in all environments. Additionally, it has yet to be shown if species diversity is constant between artificial screen measurements and all benthic substrates. It is unknown if certain species may be more likely to detach and move into the water column under certain conditions, and species have already been shown to vary in their preference of substrates (Rains & Parsons, 2015). This substrate preference also affects the traditional sampling of *Gambierdiscus* – as the choice of substrate may affect which species are detected – and if all species have a neutral preference for an artificial substrate it may enable more equal sampling of total *Gambierdiscus* diversity at a site. Irrespective of the biological implications, samples derived from artificial substrates placed in the environment for a short duration are easily comparable and much faster to process as they enable easier quantification under the microscope due to decreased detritus. This is a significant benefit to field sampling.

In the context of this study, since significant differences between sites at the same time point were found using artificial substrates, there was not complete homogenization of communities through water flow. In July there is the best comparison between artificial substrate and rubble species community composition. During that month, sites 2,3,5, and 6 all had paired screen and rubble substrate samples for which greater than 10 cells were enumerated by probes ('Bottom Lagoon', 'Control 2', 'Control 1', and 'Hot Pool' respectively). Site 5 and site 6 had similar species composition between rubble and artificial substrate methods (Figure 4.15, July). However even at these sites there remained some differences: at site 5, a small proportion of *G. belizeanus* was detected in artificial substrate samples that was not present in rubble and at site 6, a small proportion of *G. carpenteri* was present on artificial substrate but not rubble. Overall at these two locations, both substrate methods resulted in similar species composition patterns in July, as confirmed by MDS analysis (Figure 4.15, July). At sites 2 and 3, however, there were significant differences. At site 3, *G. polynesiensis* clade was present in rubble samples while this fraction was replaced largely by *G. belizeanus* and 'Unknown' in screen samples. At site 2, rubble samples showed a community largely of *G. caribaeus* and *G. polynesiensis* with a small fraction of *G. carpenteri* while screen samples were dominated by *G. australes*, *G. caribaeus*, and smaller fractions of *G. belizeanus* and 'Unknown' (Figure 4.12, Figure 4.13).

Some of these differences between rubble and screen measurements may reflect moving populations of *Gambierdiscus* in the water column. As mentioned previously, pools at Wai'Ōpae are linked by water flow largely driven by tides and northeastern trade wind-generated swell. Varying environmental conditions throughout the year likely influenced water movement at Wai'Ōpae (such as tidal strength, storms, or onshore winds), thus varying the connectivity of these sites and therefore the species composition of screen data. During relatively high tides, temperatures between pools are more similar indicating more water flow (Figure 4.5). Alternatively, an explanation for substrate community differences could be patchy distribution of species even within a site and additional replicates of rubble samples would be needed to capture true species diversity. Furthermore, the two most similar sites between rubble and screen measurements (5 and 6) also had the highest numbers of cells enumerated by probes in rubble samples from that month (50 and 25 respectively). Therefore problems with low cell numbers in samples may be driving variation between these measurements in some cases. In areas with complete connectivity, screens would represent the population of *Gambierdiscus* available to colonize substrates, and rubble samples would represent the successfully adapted species flourishing in the microenvironment.

Gambierdiscus Species Composition

Gambierdiscus species composition was examined though time, over various sites, and through different substrate methods, illustrating seasonal variation, spatial variation, and the differences between artificial substrate and rubble methods respectfully. Community composition differences seasonally may have been largely driving by increases in temperature, as mean temperatures across all pools increased from ~25°C in January to ~28.5°C in September. Furthermore, while mean temperatures at all sites roughly are within 1°C of each other at any time point, temperature variability between sites is often quite different (Figure 4.9). In January over a three day period, site 6 experienced a temperature range of more than 4°C while site 3, the most stable, only varied by~1.5°C (Figure 4.6). This range variability is captured by the temperature variance statistic (Figure 4.9). Therefore just as seasonal species differences may be driven by changes in mean temperatures, spatial differences in species composition may be driven by short-term temperature variance.

There is considerable diversity of *Gambierdiscus* species at Wai'Ōpae. There were differences in *Gambierdiscus* community composition seasonally at single sampling sites, and between sites at the same time point (Figure 4.11; Figure 4.12; Figure 4.13). In addition to those species examined using five FISH probes in this study, there are likely other species in the same environment, indicated by cells that were not labeled by any of the probes; this fraction was larger at some locations and times than others (e.g. site 6 in May, July and October). Seasonally, *Gambierdiscus australes* dominated the community at most pools and most time points. Based on the temperature gradient bar results (Figure 4.18), this likely reflects this species' ability to thrive over a wider range of temperatures than the other species examined (*G. caribaeus*).

In terms of site-specific differences, from artificial substrate data highest abundances of *Gambierdiscus* were observed at sites 3 and 5 ('Control 1' and 'Control 2'), which were most influenced by wave action and which had less variable temperatures. These sites also had the highest coral cover. Higher *Gambierdiscus* abundance could reflect favorable growth conditions due to temperature range or another environmental variable that affected growth, such as availability of macroalgae to colonize. It is interesting to note that while *G. australes* was the most common species at the high abundance sites in March and July, at those times it was much less common at sites 1 and 6 ('Hot Pool' and 'Top Lagoon') which experience a higher degree of temperature variability. Therefore, while *G. australes* can tolerate a wide range of relatively constant temperatures, it may be less suited to habitats of rapidly fluctuating temperatures. This temperature variance may have driven diversity patterns such as those seen during March, when sites 1 and 6 had the most divergent species compositions (Figure 4.14, March).

Furthermore, differences in species composition were observed between rubble and artificial substrate data (Figure 4.12; Figure 4.13). Rubble measurements overall saw less *G. australes* and more 'Unknown' cells. A higher 'Unknown' fraction could be due to rubble samples often having lower observed cell abundances, since dilution of samples took place within some samples in order to visualize *Gambierdiscus* cells among debris. However, some rubble samples of high *Gambierdiscus* abundance also contained a high percentage of the community as 'Unknown'. At sites 5 and 6 in July, high 'Unknown' fractions were present on both artificial substrate and rubble, and at high cell numbers. The community composition of these samples was very similar

overall, as supported by MDS and clustering analysis (Figure 4.15, July). In July, *G. australes* was present at five out of six sites on artificial substrate and only two out of six sites on rubble. As discussed earlier, these differences may represent distinct populations of *Gambierdiscus* that are being sampled by rubble verses artificial substrate methods, as the artificial substrate data are thought to be more sensitive to cells being transported laterally and over larger distances from other environments. In this respect, the rubble data are perhaps the most representative of the communities that adapt to each temperature regime.

Physiological Analysis of Field-Isolated Strains' Adaptation to Temperature Extremes

The temperature gradient bar analyses completed on three strains (two *G. australes* and one *G. caribaeus* strain) isolated from Wai'Ōpae gives some evidence as to why *G. caribaeus* is more abundant at higher temperatures and *G. australes* is abundant across all temperatures at Wai'Ōpae. *Gambierdiscus caribaeus* is able to grow faster at higher temperatures within the range seen at Wai'Ōpae while *G. australes* may do better at the lower temperature range. Since these growth rates are under optimum light and nutrient conditions, field growth rates are likely to be lower. The minimum temperature seen at Wai'Ōpae is at the lowest temperature for which multiple replicate cultures of *G. caribaeus* were able to grow; these conditions may thus limit growth by *G. caribaeus* during certain times of the year.

Comparing local growth rates to previously published data from other *Gambierdiscus* species from Xu et al., 2016 (Figure 4.20) and Kibler et al., 2012, *G. australes* strains were able to grow at a higher growth rate at lower temperatures than other species present at Wai'Ōpae. *Gambierdiscus polynesiensis* clade species (*G. silvae* and *G. sp type 4*) were also able to reach high growth rates at low temperatures, supporting the field observations at Wai'Ōpae (Figure 4.17). *Gambierdiscus pacificus* and *G. belizeanus* have higher growth rates at high temperatures, similar to *G. caribaeus*. At Wai'Ōpae, *G. belizeanus* was more abundant at lower temperatures. Since the Xu et al., 2016 and Kibler et al., 2012 *G. belizeanus* data were collected from strains isolated from the Caribbean, Pacific strains may have different growth rates. In Xu et al., 2016, *G. carpenteri* had consistently lower growth rates across both high and low temperatures. This supports results at Wai'Ōpae as *G. carpenteri* was the least abundant species.

Strain differences within *G. australes* are similar to differences between *Gambierdiscus* isolates observed previously in other species (Xu et al., 2016). In this analysis, the strain variation within *G. australes* was less significant than the species variation between *G. caribaeus* and *G. australes* in terms of temperature growth range since both *G. australes* strains were able to grow over similar temperatures (Figure 4.18). However in terms of amplitude of growth rates at the same temperature, the strains differed more widely. This implies that extreme temperatures may be a stronger driver of species composition changes, while strain differentiation may play more of role in fine-tuning optimal growth ranges. Further analysis of a greater number of strains is needed to establish if the strain differences observed here are due to local selection at Wai'Ōpae or merely random phenotypic variation. Stain differences observed in other studies may also be reflective of local selection.

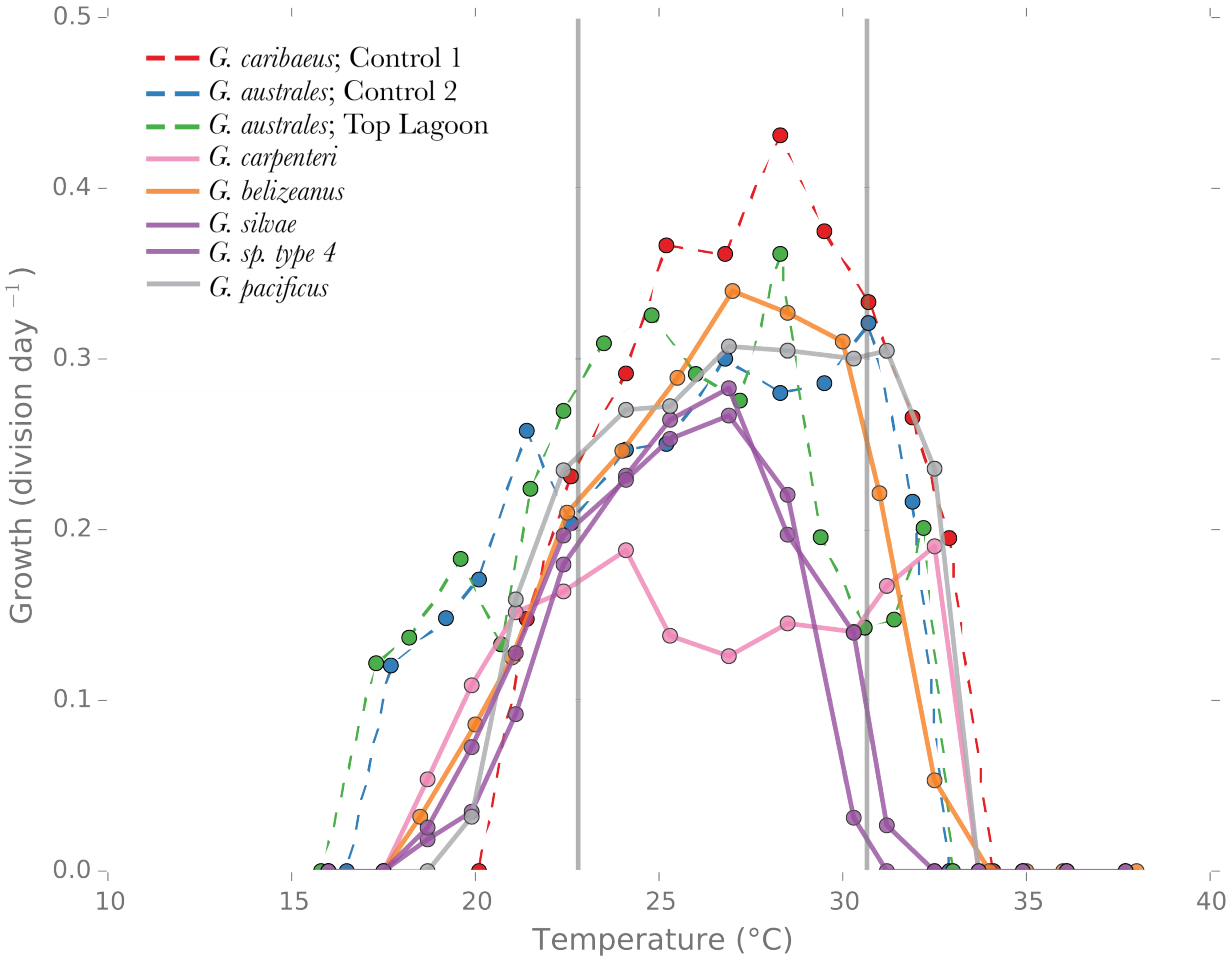


Figure 4.20 Growth Curves of *Gambierdiscus* Species

Growth curves generated from Wai'Ōpae-isolated strains (dashed lines) alongside previously published growth curves from other *Gambierdiscus* species (Xu et al., 2016). Species in purple, *G. silvae* and *G. sp. type 4*, are within the *G. polynesiensis* clade. *Gambierdiscus pacificus*, likely present at Wai'Ōpae, would be part of the 'Unknown' fraction of this study as there is no FISH probe detecting this species (grey line).

Predicted Relative Abundances of *Gambierdiscus* species from Multinomial Model

Species abundance observations were grouped by temperature (irrespective of site or month designation) and modeled over a range of five different temperatures experienced at Wai'Ōpae (25, 26, 27, and 28°C). As modeled temperatures increased from 25 to 28 °C, relative abundances of the *G. polynesiensis* clade, *G. belizeanus*, and *G. carpenteri* decreased while that of *G. caribaeus* and species represented in the 'Unknown' fraction increased. The relative abundance of *G. australes* remained high throughout but was slightly lower at 25 and 28°C than at 26 and 27°C. This supported conclusions from temperature gradient bar experiments as *G. caribaeus* had highest growth rates at higher temperatures, and *G. australes* has a larger range of temperatures permitting growth. However, it conflicted with previous growth experiments of *G. belizeanus*

which has also been shown to have high growth at high temperatures (Kibler et al., 2012, 2015; Xu et al., 2016).

If model conclusions were drawn solely from laboratory temperature growth data, at mean temperatures of 25-28°C *G. caribaeus* should dominate communities as it has a much higher growth rate over that range than other species (Figure 4.20). *G. belizeanus* has the second highest growth rate at 26 and 27°C, while the *G. australes* Top Lagoon strain has the second highest growth rate at 25 and 28°C (Figure 4.20). Moving from lower to higher mean temperatures (25 to 28°C) growth data would indicate an increase in *G. caribaeus* and *G. belizeanus* relative abundance, a slight increase in 'Unknown' due to *G. pacificus*, a decrease in the *G. polynesiensis* clade, and no change in *G. carpenteri* and average *G. australes* due to similar growth rates. This is not completely congruent with model results, and is perhaps representative of how environmental growth and community composition is more complex in the field than can be represented by laboratory measured growth rates over a single parameter, though these rates can be informative towards explaining environmental patterns of community composition. Since this model only takes into account mean temperatures across Wai'Ōpae, it largely interprets seasonal trends. Further examination of temperature variance data would be necessary to fully capture site heterogeneity.

The Role of Ciguatera in Deteriorating Reef Systems

The sampling for this study took place at an unusual period of time for Wai'Ōpae, Hawai'i. In 2014 Hurricane Iselle caused significant damage to the reef and surrounding areas and in 2015, record hot temperatures over the summer further stressed corals. Many corals were broken by the hurricane and further individuals were at least partially bleached the following summer. *Gambierdiscus* generally increases in abundance as coral reefs decline in health, likely due to opportunistic increases in its preferred macroalgal substrates. The hurricane-induced damage and nutrient influx from surrounding homes could have stimulated overall macroalgal growth and set off a successional cascade for macroalgae with new broken coral rubble for colonization. Interestingly, this study instead found the highest *Gambierdiscus* abundances in the most coral dominated areas of Wai'Ōpae, sites 3 and 5. Turf algae was abundant across most sites but site 6 was the only site with large amounts of macroalgae and consistently had lower *Gambierdiscus* abundances. However, if cells are attaching to artificial substrates mostly though water flowing across the screen, then this would also explain why areas exposed to more wave action would also appear to have higher *Gambierdiscus* abundance.

Conclusion

Wai'Ōpae, Hawai'i is a diverse site both in its environmental drivers and its *Gambierdiscus* community. The intent of this study was to infer general trends that may hold true for global *Gambierdiscus* communities as seawater temperatures increase due to climate change. Within Wai'Ōpae, this study found quickly shifting communities that may be responding to temperature fluctuations across the site. *Gambierdiscus* abundance is known to be patchy and this study reinforced these findings. Across a relatively small geographic area, there were hot spots of abundance of different species. Site 6, 'Hot Pool', appeared to be the most different from other sites, with the highest temperature variance potentially affecting its often-unique species composition. Furthermore, the furthest seaward sites, sites 3 and 5, consistently had the highest

Gambierdiscus genus abundance based on artificial substrate measurements whenever data were available. These sites were also the most thermally stable. High diversity in *Gambierdiscus* abundance and community composition were observed spatially and seasonally. January had the lowest total *Gambierdiscus* abundance (aside from sites 3 and 5, for which data are incomplete).

Gambierdiscus abundance and growth rate has been shown to be positively correlated with higher sea surface temperatures within its range of temperature tolerance (Dickey & Plakas, 2010; Llewellyn, 2010; Parsons, Settlemier, & Bienfang, 2010), and a hypothesis has been raised that this will result in increased ciguatoxicity. The Wai'Ōpae data support this correlation as January had lower mean seawater temperatures across sites, and the lowest *Gambierdiscus* genus abundance. However, periodic higher temperatures locally, such as in site 1 'Top Lagoon', did not translate to higher *Gambierdiscus* abundances, but instead a shift in species composition and lower total abundance.

Interestingly, evidence from the Gulf of Mexico and Caribbean indicates that species that are currently expanding their ranges may not be the most toxic members of the *Gambierdiscus* genus. In the Flower Garden Banks National Marine Sanctuary in the northern Gulf of Mexico as well as on the eastern coast of the United States, *G. carolineanus* is prevalent while members of the *G. polynesiensis* clade, which are currently thought to be the most ciguatoxic species, are absent (P. A. Tester et al., 2010; P. a. Tester et al., 2013). However, this study found that members of the *G. polynesiensis* clade had higher relative abundance at lower mean temperatures, perhaps indicating their ability to expand their range to more northern and southern latitudes, at least in the Pacific Ocean. Furthermore, there is the added complexity that *Gambierdiscus* species which are considered non-toxic in terms of ciguatoxin may actually still produce compounds that contribute to the poisoning syndrome. Since CFP involves the biotransformation of gambiertoxins to ciguatoxins within fish, compounds that are not considered toxic in their algal form may be biotransformed to form part of the toxic assemblage. Further fish feeding studies using various species of *Gambierdiscus* would help discriminate these effects.

Overall, this study found complex dynamics at play between environmental conditions and *Gambierdiscus* community composition that illustrate the difficulty of predicting future risk of ciguatera fish poisoning amid climate change conditions. Multinomial modeling or species relative abundance counts resulted in: 1) predicted higher relative abundance of *G. caribaeus* at higher temperatures; 2) higher relative abundance of *G. belizeanus*, *G. carpenteri*, and *G. polynesiensis* clade at lower temperatures; and 3) high relative abundance of *G. australes* across all temperatures. Temperature growth curves for two *G. australes* strains and one *G. caribaeus* strain support this result. However, low abundance of *Gambierdiscus* at thermally variable sites as well as altered community composition compared to control sites would suggest that as coastal seawater temperatures warm and likely become more variable with global warming, *Gambierdiscus* populations may decline in abundance and switch from *G. australes* domination to *G. caribaeus*, *G. polynesiensis clade*, *G. carpenteri* and *G. belizeanus* dominance, as observed at site 6. Finally, sampling bias between rubble and artificial sampling methods is a critical factor that needs further investigation in the context of these types of studies. 24-hour screen deployments may be convenient, but also may not reveal true adaptation of the *Gambierdiscus* species to microenvironments. The use of FISH probes to differentiate species abundance, however, enables further studies on environmental effects on community composition and when used in a

monitoring capacity for CFP risk, allows reactionary measures to be taken if toxic species increase in abundance.

Funding Sources:
USFDA (F223201000060C)
NOAA NOS (CiguaHAB; NA11NOS4780060 and NA11NOS4780028);
COHH - Woods Hole Center for Oceans and Human Health – NSF Grant OCE-1314642 and NIEHS 1P01-ES021923-01
WHOI Ocean Ventures Fund; WHOI Coastal Institute Research Proposal Award

Chapter 5

Conclusion

Thesis Summary and Future Directions

Multiple dinoflagellates produce toxins and are responsible for poisoning syndromes in humans. Two genera of these harmful algal bloom (HAB)-forming species are *Alexandrium* and *Gambierdiscus*. The mechanisms of their bloom formation and accumulation of toxins are fundamentally different, the former being a planktonic species whose toxins directly accumulate in shellfish while the later is a benthic species whose toxin bioaccumulates through tropical food web transfer. Nevertheless, the methods used to study the genetic basis for their interspecific variation are similar. Dinoflagellates have long been a genetically intractable group, and the application of transcriptome analysis to these organisms, as described here, is an important step in understanding the hidden genetic diversity that underlies physiological differences between species. In both *Alexandrium* and *Gambierdiscus*, species exhibit marked differences in toxicity and growth under varying conditions. This thesis first aimed to illustrate the genetic basis for saxitoxin production in *Alexandrium* (Chapter 2), and establish further evidence for a Horizontal Gene Transfer (HGT) event that brought the saxitoxin biosynthesis pathway from cyanobacteria to dinoflagellates. The genus *Alexandrium* contains many closely related toxic and non-toxic species, which, analyzed together and with other saxitoxin-producing dinoflagellates from the genera *Pyrodinium* and *Gymnodinium*, allowed the identification of differences in saxitoxin gene content between toxic and non-toxic species. This analysis illustrated genetic changes that may be important for toxin biosynthesis and further elucidated the evolutionary history of the saxitoxin gene cluster in dinoflagellates. Furthermore, the discovery of the retention and expression of many saxitoxin biosynthesis genes in a non-saxitoxin producing genus, *Gambierdiscus*, showed the capability of dinoflagellates to 'recycle' genes from one pathway to another. In addition, this study confirmed previous reports of the requirement of the domain sxtA4 and a clade of sxtG for saxitoxin production, as both transcripts were only found in toxic species. Surprisingly, along with an increased diversity in sxtA transcripts, this study also found clusters of *Gambierdiscus* transcripts from multiple *Gambierdiscus* species paired with several of these sxtA clades. Overall, Chapter 2 illustrates the genetic diversity that occurs through gene duplication and diversification of a biosynthesis pathway in dinoflagellates.

Gambierdiscus, like *Alexandrium*, has both toxic and non-toxic member species. Due to *Gambierdiscus* species' cryptic diversity and variance in toxicity, species-specific enumeration is an important component in determining regional risk of ciguatera fish poisoning (CFP). The development of species-specific fluorescent *in situ* hybridization (FISH) probes allowed analysis of field samples to determine *Gambierdiscus* community composition (Chapter 3). This novel approach allowed the quantitative assessment of species along a six-month seasonal increase in *Gambierdiscus* abundance in Heine Grassbed (HGB), Florida Keys. A shift in *Gambierdiscus* community composition occurred as total cell abundance increased: the relative abundance of *G. caribaeus* and *G. carpenteri* increased while that of the *G. polynesiensis* clade decreased. Such shifts in species abundance have potential implications for risk of ciguatera fish poisoning as an increase in abundance of *Gambierdiscus* which was due largely to a non-toxic species would not increase overall risk of CFP in that area.

One important area of current concern is that shifts in *Gambierdiscus* community composition and abundance may occur with warming seawater temperatures. Since *Gambierdiscus* species have different optimal ranges of temperature for growth, changing seawater temperatures may cause species ranges to expand or shrink, affecting community composition and local risk of CFP. In

Chapter 4, novel FISH probes were applied to a thermally variable environment in Hawai'i. Here, phenotypic diversity in temperature tolerance between species was shown to affect the diverse *Gambierdiscus* community, as seasonally driven increases in mean temperatures and spatially variable temperature fluctuations caused shifts in the relative abundance of *Gambierdiscus* species. *Gambierdiscus* is known to be patchy in distribution and abundance, and this study reinforced those findings. Across a relatively small geographic area, there were hot spots of abundance of different species. The most thermally stable sites had the highest overall *Gambierdiscus* abundance, while more thermally variable sites closer to shore often had lower abundance and divergent species compositions. Multinomial modeling of species relative abundances resulted in predicted: 1) higher relative abundance of *G. caribaeus* at higher temperatures; 2) higher relative abundance of *G. belizeanus*, *G. carpenteri*, and *G. polynesiensis* clade at lower temperatures; and 3) high relative abundance of *G. australes* across all temperatures. Temperature growth analysis of two *G. australes* strains and one *G. caribaeus* strain supported this result as *G. caribaeus* has higher growth rates at high temperatures while *G. australes* was able to grow across a wider range of temperatures. Overall, this study found complex dynamics at play between environmental conditions and *Gambierdiscus* community composition that illustrate the difficulty of predicting future risk of ciguatera fish poisoning amid climate change conditions. However, one conclusion from this analysis is the importance of thermal variability on *Gambierdiscus* community composition. Although all sites had mean temperatures within ~1°C of each other, they experienced different ranges in temperature which dramatically affected species composition. If under future climate change scenarios, near-shore temperatures are expected both to rise and be more variable, this may have very different consequences on *Gambierdiscus* species abundance than an increase in temperatures alone.

Dinoflagellates occupy diverse roles in marine ecosystems due to both their mixotrophic lifestyle, (which is becoming increasingly important as a major mode of nutrition for photosynthetic dinoflagellates: Jeong et al., 2010), and their ability to live in a diverse array of environments from pelagic to coastal to benthos. Furthermore, the ability of some species to produce toxins allows them to affect multiple trophic levels in a manner similar to how a keystone species exerts a disproportionally large influence on an ecosystem (Zimmer & Ferrer, 2007). Coupled with their unusual cellular molecular biology and high rates of HGT, substantial and important research questions remain to be answered about their ecology, genetics, and physiology. Overall, this thesis strived to use the ecologically important dinoflagellate genera of *Gambierdiscus* and *Alexandrium* to illustrate the genetic variability within diverse dinoflagellate genera, how this variability results in phenotypic variability, and how phenotypic traits determine environmental community composition. Many studies remain to be done: some more methodological, such as a closer examination of the differences in substrate use for *Gambierdiscus* sampling, and some more broad, such as further elucidating the influence of HGT events on toxic dinoflagellates. The discussion around choice of substrate for *Gambierdiscus* populations has been focused on comparing total *Gambierdiscus* abundance through different methods. Potential species composition differences between methods should also be examined since this has particular significance towards the ecology of *Gambierdiscus* and estimation of risk of CFP. In terms of HGT events, only recently has the significance of these processes within eukaryotic microbial populations been emphasized. This process seems to be especially prevalent in dinoflagellates and has resulted in important ecological consequences for the group. As more sequencing data become available for dinoflagellates, especially diverse and ecologically important genera such as *Alexandrium* and *Gambierdiscus*, examination of rates and presence of HGT events will be

informative as these events affect phenotypic traits which can have important ramifications for human health.

The consequences of globally warming seawater temperatures on *Gambierdiscus* populations are of special concern due to changing risk of CFP. With respect to determination of temperature tolerance, further experiments in regions of natural temperature gradients that are not as closely linked as at Wai'Ōpae Tide Pools, Hawai'i, would allow the greater separation of adapted populations of *Gambierdiscus* cells from those present due to water transportation across the site. Furthermore, in addition to a greater number of strains analyzed under the temperature gradient bar to determine plasticity of temperature response within a single species, temperature experiments incorporating temperature variability would be useful. Such experiments could support findings of this analysis that temperature variability may represent a different stressor to *Gambierdiscus* populations than high mean temperature alone. In the future, this thesis presents developments in our ability to distinguish between dinoflagellates, both genetically and within environmental samples, that will enable further investigations into the ecology of these extraordinarily complex and important species.

Chapter 2
Supplemental Information

Target	Origin	Gene	Species
Actin	gi|134037068|gb|ABO47874.1	actin	Alexandrium fundyense
Chlabp	gi|62638121|gb|AAX92666.1	chloroplast peridinin-chlorophyll a-binding protein precursor	Alexandrium tamarense
Luciferase	gi|54610717|gb|AAV35378.1	luciferase	Alexandrium tamarense
Luciferinbp	gi|394309363|gb|AFN27012.1	luciferin binding protein, partial	Alexandrium tamarense
Mitcyc1	gi|53760455|gb|AAU93350.1	mitotic cyclin 1	Lingulodinium polyedrum
PCNA	gi|133856052|gb|ABO40135.1	proliferating cell nuclear antigen	Alexandrium fundyense
Rubisco	gi|37727276|gb|AAO13079.1	ribulose 1,5-bisphosphate carboxylase oxygenase form II, partial	Prorocentrum minimum
SxtA	gi|328926200|gb|ADY62525.1	SxtA long isoform precursor	Alexandrium fundyense
SxtA	gi|328926199|gb|ADY62524.1	SxtA short isoform precursor	Alexandrium fundyense
SxtA	Hackett et al, 2012	Atamclcbn58851	Alexandrium pacificum
SxtA	Hackett et al, 2012	Atamclcbn93306	Alexandrium pacificum
SxtA	Hackett et al, 2012	Atamclcbn86513	Alexandrium pacificum
SxtA	Hackett et al, 2012	Atamclcbn31894	Alexandrium pacificum
SxtA	Hackett et al, 2012	Atam98411_1_550_1	Alexandrium pacificum
SxtA	Hackett et al, 2012	Atam98432_1_639_2	Alexandrium pacificum
SxtA	gi|114462387|gb|ABI75123.1	polyketide synthase-related protein	Dolichospermum circinale AWQC131C
SxtA	gi|114462352|gb|ABI75094.1	polyketide synthase-related protein	Cylindrospermopsis raciborskii T3
SxtA	gi|282897235|ref|ZP_06305237.1	sxtA	Raphidiopsis brookii D9
SxtA	gi|195984481|gb|ACG63826.1	SxtA	Lyngbya wollei
SxtA	gi|195984455|gb|ACG63801.1	SxtA	Aphanizomenon sp. NH-5
SxtB	Hackett et al, 2012	Atam22328_455_1327_2	Alexandrium pacificum
SxtB	Hackett et al, 2012	Atam7422_66_1034_3	Alexandrium pacificum
SxtB	gi|114462386|gb|ABI75122.1	cytidine deaminase	Dolichospermum circinale AWQC131C
SxtB	gi|195984482|gb|ACG63827.1	SxtB	Lyngbya wollei
SxtB	gi|195984454|gb|ACG63800.1	SxtB	Aphanizomenon sp. NH-5
SxtB	gi|282897236|ref|ZP_06305238.1	sxtB (Cytidine deaminase)	Raphidiopsis brookii D9
SxtB	gi|114462351|gb|ABI75093.1	cytidine deaminase	Cylindrospermopsis raciborskii T3
SxtC	gi|282897237|ref|ZP_06305239.1	hypothetical protein CRD_02159	Raphidiopsis brookii D9
SxtC	gi|269114765|gb|ACZ26227.1	SxtC	Lyngbya wollei
SxtC	gi|269114761|gb|ACZ26224.1	SxtC	Aphanizomenon sp. NH-5
SxtC	gi|114462350|gb|ABI75092.1	SxtC	Cylindrospermopsis raciborskii T3
SxtD	gi|195984502|gb|ACG63847.1	SxtD	Lyngbya wollei
SxtD	gi|195984458|gb|ACG63804.1	SxtD	Aphanizomenon sp. NH-5
SxtD	gi|114462347|gb|ABI75089.1	sterole desaturase	Cylindrospermopsis raciborskii T3
SxtD	gi|282897239|ref|ZP_06305241.1	sxtD (Sterol desaturase-like protein)	Raphidiopsis brookii D9
SxtD	gi|114462389|gb|ABI75125.1	sterole desaturase	Anabaena circinalis AWQC131C
SxtE	gi|269114766|gb|ACZ26228.1	SxtE	Lyngbya wollei
SxtE	gi|269114762|gb|ACZ26225.1	SxtE	Aphanizomenon sp. NH-5
SxtE	gi|114462353|gb|ABI75095.1	SxtE	Cylindrospermopsis raciborskii T3
SxtE	gi|114462388|gb|ABI75124.1	SxtE	Anabaena circinalis AWQC131C
SxtE	gi|282897234|ref|ZP_06305236.1	sxtE	Raphidiopsis brookii D9
SxtF	gi|114462354|gb|ABI75096.1	sodium-driven multidrug and toxic compound extrusion protein	Cylindrospermopsis raciborskii T3
SxtF	gi|282897233|ref|ZP_06305235.1	sxtF (NorM)	Raphidiopsis brookii D9
SxtG	gi|443418936|gb|AGC84341.1	amidinotransferase	Alexandrium minutum
SxtG	gi|443418932|gb|AGC84339.1	amidinotransferase	Alexandrium fundyense
SxtG	Hackett et al, 2012	Atam8098_299_1231_1	Alexandrium pacificum
SxtG	Hackett et al, 2012	Atam87049_31_621_1	Alexandrium pacificum
SxtG	gi|195984485|gb|ACG63830.1	SxtG	Lyngbya wollei
SxtG	gi|195984467|gb|ACG63813.1	SxtG	Aphanizomenon sp. NH-5
SxtG	gi|194740205|gb|ACF94640.1	amidinotransferase	Cylindrospermopsis raciborskii T3
SxtG	gi|282897232|ref|ZP_06305234.1	sxtG (Amidinotransferase)	Raphidiopsis brookii D9
SxtG	gi|194740215|gb|ACF94645.1	amidinotransferase	Anabaena circinalis AWQC131C
SxtH	gi|195984486|gb|ACG63831.1	SxtH	Lyngbya wollei
SxtH	gi|195984468|gb|ACG63814.1	SxtH	Aphanizomenon sp. NH-5
SxtH	gi|194740229|gb|ACF94652.1	dioxygenase	Anabaena circinalis AWQC173A
SxtH	gi|194740221|gb|ACF94648.1	dioxygenase	Cylindrospermopsis raciborskii T3
SxtH	gi|282897231|ref|ZP_06305233.1	sxtH (Rieske (2Fe-2S) region protein)	Raphidiopsis brookii D9
SxtI	gi|195984470|gb|ACG63816.1	SxtI	Aphanizomenon sp. NH-5
SxtI	gi|269114767|gb|ACZ26229.1	truncated SxtI	Lyngbya wollei
SxtI	gi|282897230|ref|ZP_06305232.1	sxtI (Putative carbamoyltransferase)	Raphidiopsis brookii D9
SxtI	gi|114462357|gb|ABI75099.1	NodU/CmcH-related carbamoyltransferase	Cylindrospermopsis raciborskii T3
SxtI	gi|114462403|gb|ABI75139.1	NodU/CmcH-related carbamoyltransferase	Anabaena circinalis AWQC131C

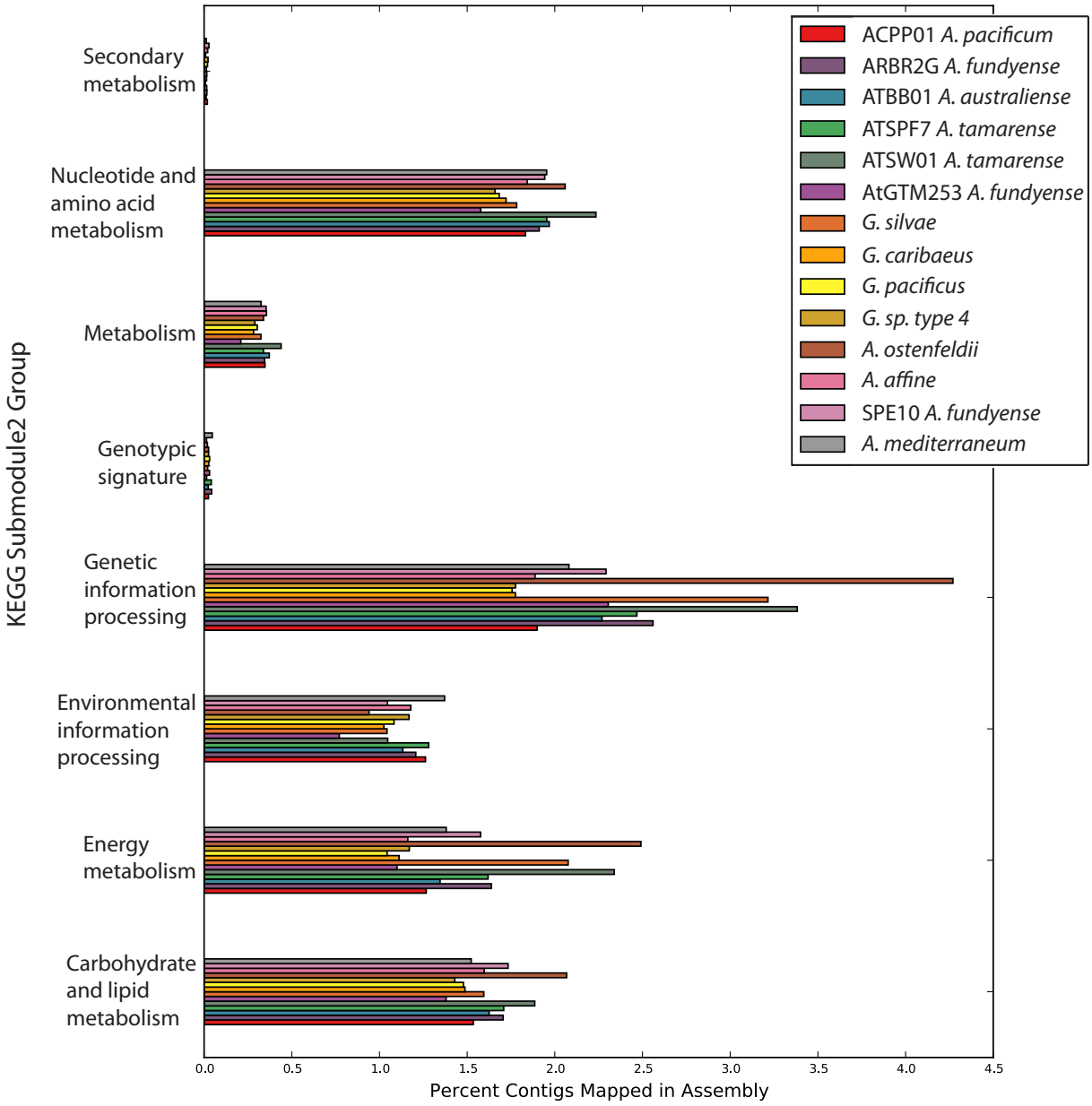
Supplemental Table 2.1 Queries used in BLAST Analysis (Part 1)

List of queries, their GeneBank accession numbers or study of origin, gene and species designations.

Target	Origin	Gene	Species
SxtJ	gi|195984471|gb|ACG63817.1	SxtJ	Aphanizomenon sp. NH-5
SxtJ	gi|114462358|gb|ABI75100.1	SxtJ	Cylindrospermopsis raciborskii T3
SxtJ	gi|114462404|gb|ABI75140.1	SxtJ	Anabaena circinalis AWQC131C
SxtJ	gi|282897229|ref|ZP_06305231.1	sxtJ	Raphidiopsis brookii D9
SxtK	gi|195984472|gb|ACG63818.1	SxtK	Aphanizomenon sp. NH-5
SxtK	gi|114462359|gb|ABI75101.1	SxtK	Cylindrospermopsis raciborskii T3
SxtK	gi|114462405|gb|ABI75141.1	SxtK	Anabaena circinalis AWQC131C
SxtL	gi|195984473|gb|ACG63819.1	SxtL	Aphanizomenon sp. NH-5
SxtL	gi|282897228|ref|ZP_06305230.1	sxtL	Raphidiopsis brookii D9
SxtL	gi|114462360|gb|ABI75102.1	GDSL-lipase	Cylindrospermopsis raciborskii T3
SxtL	gi|114462406|gb|ABI75142.1	GDSL-lipase	Anabaena circinalis AWQC131C
SxtM	gi|195984469|gb|ACG63815.1	SxtM	Aphanizomenon sp. NH-5
SxtM	gi|195934015|gb|ACG58379.1	MATE transporter	Anabaena circinalis 131C
SxtM	gi|114462361|gb|ABI75103.1	sodium-driven multidrug and toxic compound extrusion protein	Cylindrospermopsis raciborskii T3
SxtM	gi|282897225|ref|ZP_06305227.1	sxtM (NorM)	Raphidiopsis brookii D9
SxtN	gi|195984466|gb|ACG63812.1	SxtN	Aphanizomenon sp. NH-5
SxtN	gi|114462362|gb|ABI75104.1	sulfotransferase	Cylindrospermopsis raciborskii T3
SxtN	gi|114462399|gb|ABI75135.1	sulfotransferase	Anabaena circinalis AWQC131C
SxtN	gi|282897227|ref|ZP_06305229.1	sxtN (Putative sulfotransferase protein)	Raphidiopsis brookii D9
SxtO	gi|114462407|gb|ABI75143.1	adenylylsulfate kinase	Dolichospermum circinale AWQC131C
SxtO	gi|114462373|gb|ABI75115.1	adenylylsulfate kinase	Cylindrospermopsis raciborskii T3
SxtP	gi|195984501|gb|ACG63846.1	SxtP	Lyngbya wollei
SxtP	gi|195984459|gb|ACG63805.1	SxtP	Aphanizomenon sp. NH-5
SxtP	gi|114462372|gb|ABI75114.1	putative saxitoxin-binding protein	Cylindrospermopsis raciborskii T3
SxtP	gi|114462390|gb|ABI75126.1	putative saxitoxin-binding protein	Anabaena circinalis AWQC131C
SxtP	gi|282897219|ref|ZP_06305221.1	sxtP (Integrins alpha chain)	Raphidiopsis brookii D9
SxtQ	gi|195984499|gb|ACG63844.1	SxtQ	Lyngbya wollei
SxtQ	gi|195984460|gb|ACG63806.1	SxtQ	Aphanizomenon sp. NH-5
SxtQ	gi|114462371|gb|ABI75113.1	SxtQ	Cylindrospermopsis raciborskii T3
SxtQ	gi|282897220|ref|ZP_06305222.1	sxtQ	Raphidiopsis brookii D9
SxtR	gi|195984498|gb|ACG63843.1	SxtR	Lyngbya wollei
SxtR	gi|195984461|gb|ACG63807.1	SxtR	Aphanizomenon sp. NH-5
SxtR	gi|282897221|ref|ZP_06305223.1	sxtR	Raphidiopsis brookii D9
SxtR	gi|114462370|gb|ABI75112.1	acyl-CoA N-acyltransferase	Cylindrospermopsis raciborskii T3
SxtR	gi|114462392|gb|ABI75128.1	acyl-CoA N-acyltransferase	Anabaena circinalis AWQC131C
SxtS	gi|269114763|gb|ACG63841.2	SxtS	Lyngbya wollei
SxtS	gi|195984463|gb|ACG63809.1	SxtS	Aphanizomenon sp. NH-5
SxtS	gi|114462368|gb|ABI75110.1	phytanoyl-CoA dioxygenase	Cylindrospermopsis raciborskii T3
SxtS	gi|114462396|gb|ABI75132.1	phytanoyl-CoA dioxygenase	Anabaena circinalis AWQC131C
SxtS	gi|282897222|ref|ZP_06305224.1	sxtS (Phytanoyl-CoA dioxygenase family protein)	Raphidiopsis brookii D9
SxtT	gi|195984495|gb|ACG63840.1	SxtT	Lyngbya wollei
SxtT	gi|195984464|gb|ACG63810.1	SxtT	Aphanizomenon sp. NH-5
SxtT	gi|114462367|gb|ABI75109.1	phenylpropionate dioxygenase	Cylindrospermopsis raciborskii T3
SxtT	gi|114462397|gb|ABI75133.1	phenylpropionate dioxygenase	Anabaena circinalis AWQC131C
SxtT	gi|282897223|ref|ZP_06305225.1	sxtT (Rieske (2Fe-2S) region protein)	Raphidiopsis brookii D9
SxtU	gi|195984494|gb|ACG63839.1	SxtU	Lyngbya wollei
SxtU	gi|195984465|gb|ACG63811.1	SxtU	Aphanizomenon sp. NH-5
SxtU	gi|114462366|gb|ABI75108.1	short-chain alcohol dehydrogenase	Cylindrospermopsis raciborskii T3
SxtU	gi|114462398|gb|ABI75134.1	short-chain alcohol dehydrogenase	Anabaena circinalis AWQC131C
SxtU	gi|282897224|ref|ZP_06305226.1	sxtU (Short-chain dehydrogenase/reductase SDR)	Raphidiopsis brookii D9
SxtV	gi|195984492|gb|ACG63837.1	SxtV	Lyngbya wollei
SxtV	gi|114462365|gb|ABI75107.1	FAD-dependent succinate dehydrogenase/fumarate reductase	Cylindrospermopsis raciborskii T3
SxtW	gi|269114768|gb|ACZ26230.1	SxtW	Lyngbya wollei
SxtW	gi|195984456|gb|ACG63802.1	SxtW	Aphanizomenon sp. NH-5
SxtW	gi|114462364|gb|ABI75106.1	ferredoxin	Cylindrospermopsis raciborskii T3
SxtX	gi|195984491|gb|ACG63836.1	SxtX	Lyngbya wollei
SxtX	gi|195984457|gb|ACG63803.1	SxtX	Aphanizomenon sp. NH-5
SxtX	gi|194740237|gb|ACF94656.1	saxitoxin N1-hydroxylase	Cylindrospermopsis raciborskii T3
SxtY	gi|114462375|gb|ABI75117.1	phosphate uptake regulator	Cylindrospermopsis raciborskii T3
SxtZ	gi|114462376|gb|ABI75118.1	histidine kinase	Cylindrospermopsis raciborskii T3

Supplemental Table 2.1 Queries used in BLAST Analysis (Part 2)

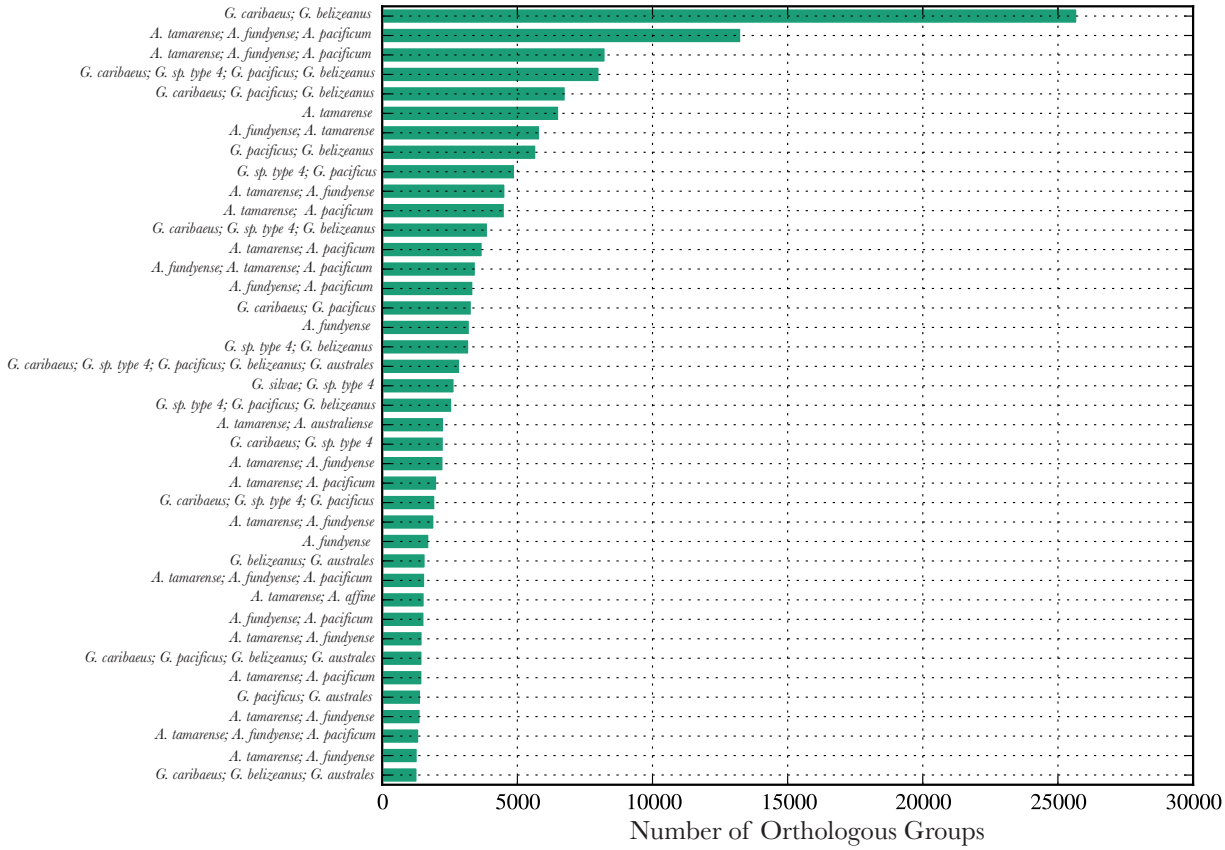
List of queries, their GeneBank accession numbers or study of origin, gene and species designations.



Supplemental Figure 2.1: KEGG Analysis of Transcriptome Libraries
To translate contigs into their most likely protein sequence open reading frame sequences of the contig libraries were predicted and translated using *OrfPredictor* (Min et al., 2005). As input a blastx results file of the contig library verses NCBI's nr database was provided in addition to the library's sequence fasta file to improve prediction accuracy. The resulting output of predicted amino acid sequences were then fed into the *Kyoto Encyclopedia of Genes and Genomes'* (KEGG) online Ghost Koala interface (Kanehisa et al., 2016). This analysis was completed for fourteen transcriptomes in total, at least one from each species of the *A. tamarense* species complex. This figure shows the percent of contigs in an assembly which mapped to a known protein function contained within these large submodules.

Orthologous Sequence Identification

In order to identify orthologs between the translated transcriptomes and assess transcriptome similarity between species, OrfPredictor outputs of predicted amino acid sequences from each were analyzed with proteinortho, a program that uses graph clustering and BLAST similarity scores to discriminate between orthologous and paralogous sequences. For each library that was translated through *OrfPredictor*, the protein fasta file was analyzed with *proteinortho* (Lechner et al., 2011) to group the contigs into orthologous groups of sequences between at least two libraries.



Supplemental Figure 2.2 Orthologous Groups Shared Between Transcriptome Libraries

This figure illustrates the size of the largest 40 groups of orthologous sequences shared between transcriptome libraries predicted by the software package *proteinortho*. *Gambierdiscus caribaeus* and *G. belizeanus* share the most groups between them. When species pairs repeat, it is due to different strains of the same species contained within those pairs.

Chapter 4
Supplemental Information

Site	p-value
1	1.00
2	0.83
3	0.97
4	0.95
5	0.31
6	0.17

Supplemental Table 4.1 ANOVA Results by Month

One-way ANOVA testing was use to compare overall *Gambierdiscus* abundance between 24 and 72 hour screen deployments. Individual hybridization abundance measurements were grouped within each site. Null hypothesis could not be rejected of equal mean abundance of 24 and 72 hour screen deployments.

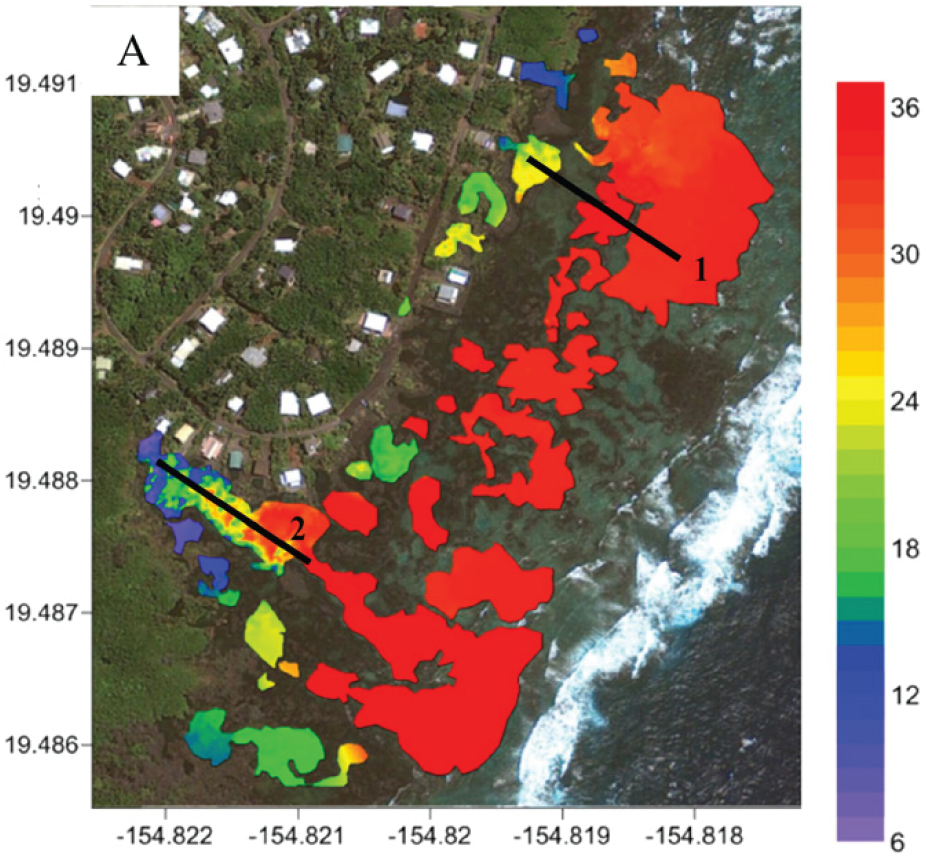
Site	Month	p-value
1	January	0.61
	March	0.80
	July	0.61
2	January	0.70
	March	0.92
	July	0.38
3	January	0.08
	March	0.41
	July	0.75
4	January	1.00
	March	0.53
	July	0.48
5	January	0.62
	March	0.37
	July	0.98
6	January	0.67
	March	0.35
	July	0.44

Supplemental Table 4.2 ANOVA Results by Site and Month

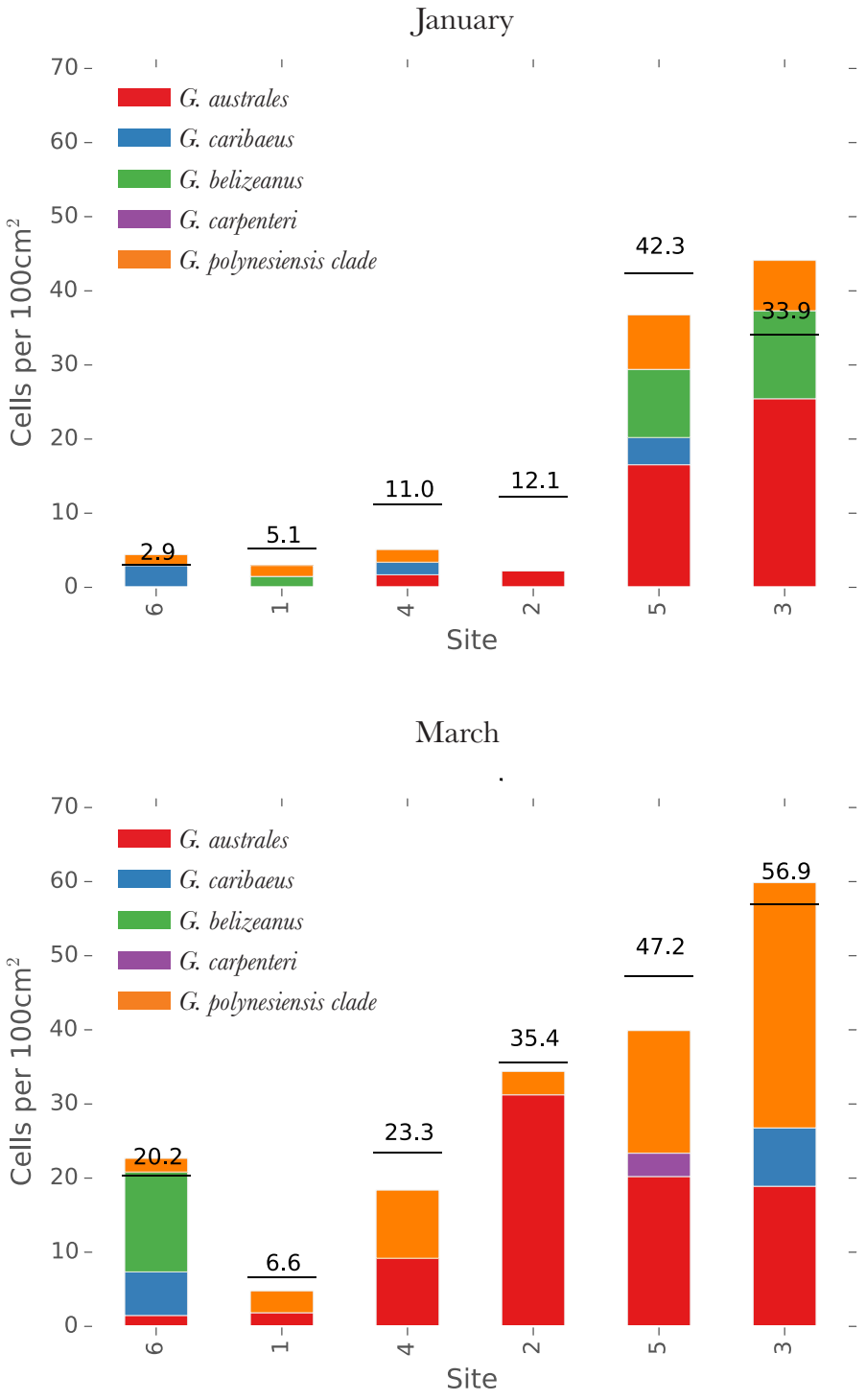
One-way ANOVA testing was use to compare overall *Gambierdiscus* abundance between 24 and 72 hour screen deployments. Individual hybridization abundance measurements were grouped within each site and month. Null hypothesis could not be rejected of equal mean abundance of 24 and 72 hour screen deployments.

Salinity		
Site	4/1/15	7/16/15
1	35.0	29.0
2	35.0	32.1
3	35.0	34.5
4	35.0	32.5
5	35.0	30.0
6	35.0	32.0

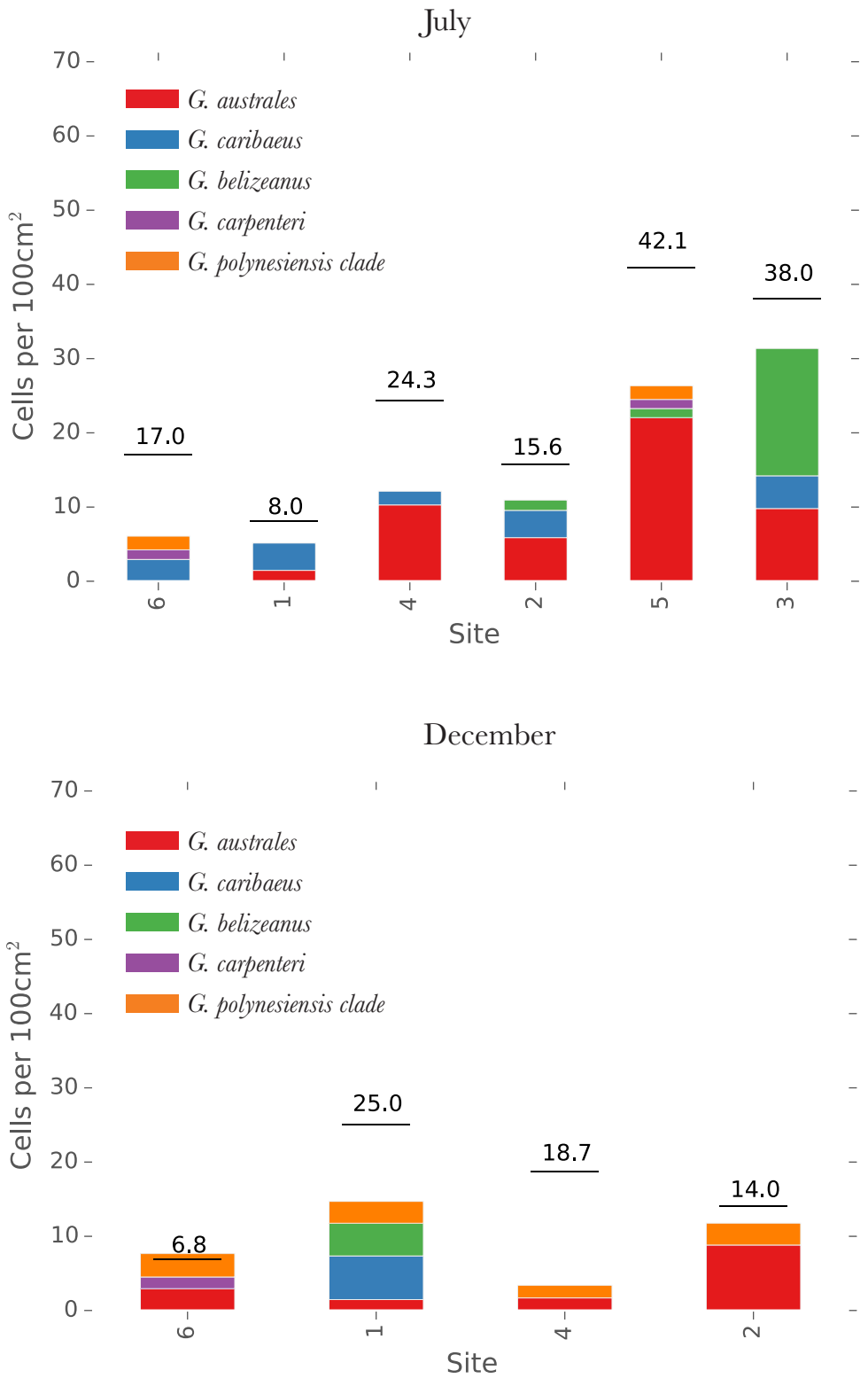
Supplemental Table 4.3 Salinity values at Wai'Ōpae
Table of salinity values taken at depth on 4/1/2015 and 7/16/2015 by refractometer.



Supplemental Figure 4.1 (From from Wiegner et al., 2016)
Salinity and transect designation for nutrient study (Wiegner et al., 2016). Transect 1 had three times higher nutrient values than transect 2 for dissolved NO3− + NO2− and TDN. Three of the sites in this study are along 'Transect 2' in the lagoon. TDN ranged from 18 to 11 μmol L^{-1} across Transect 2. (0m to 100m from shore). The furthest site in this study is along this transect, Site 5 'Control 1', and is 234m from shore.

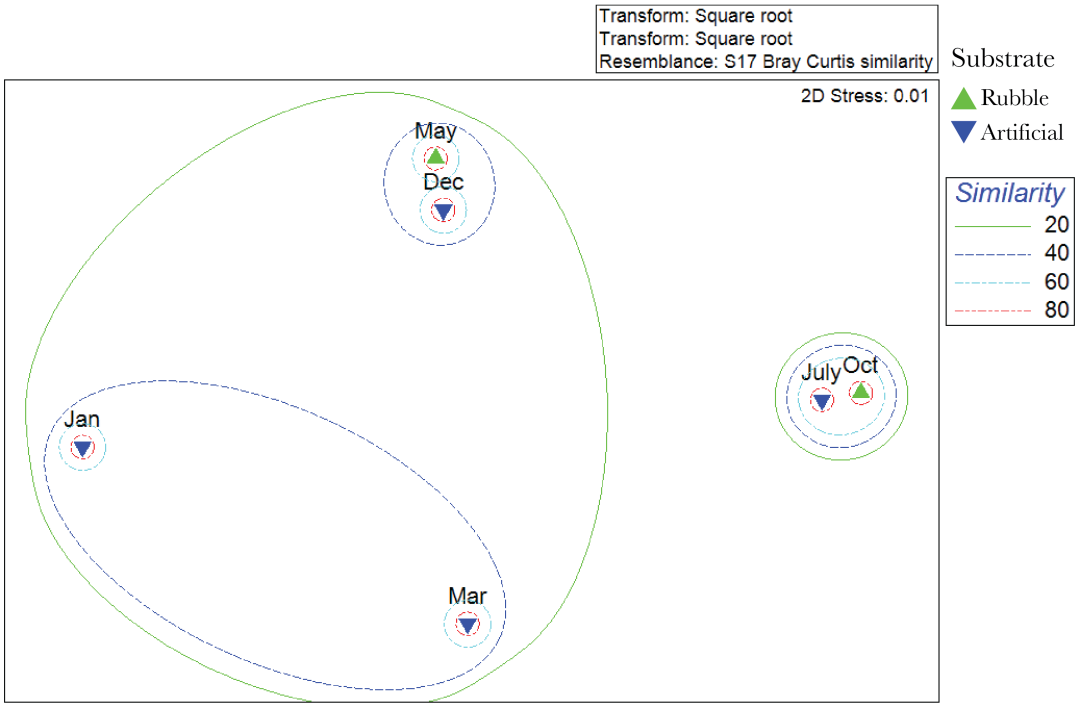


Supplemental Figures (4.2-4.3)
Individual months plotted with abundance of each *Gambierdiscus* species and total abundance of *Gambierdiscus* as a genus per 100cm² of artificial substrate screen area.

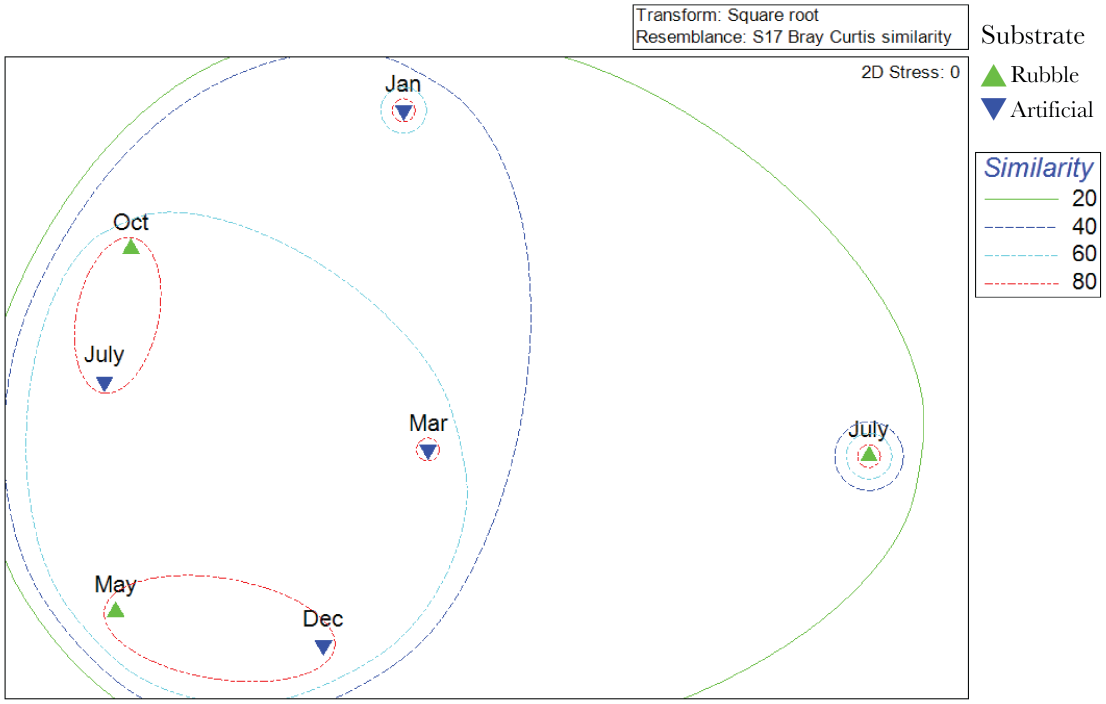


Supplemental Figures (4.4-4.5)
Individual months plotted with abundance of each *Gambierdiscus* species and total abundance of *Gambierdiscus* as a genus per 100cm^2 of artificial substrate screen area.

Site 1 Top Lagoon



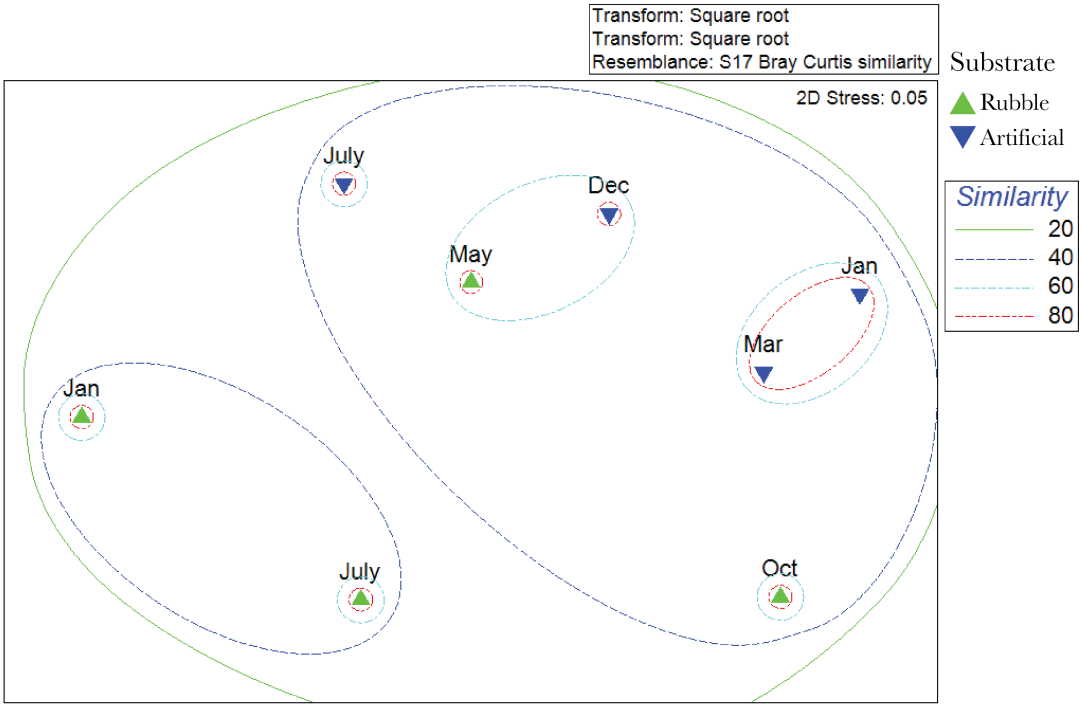
Site 4 Middle Lagoon



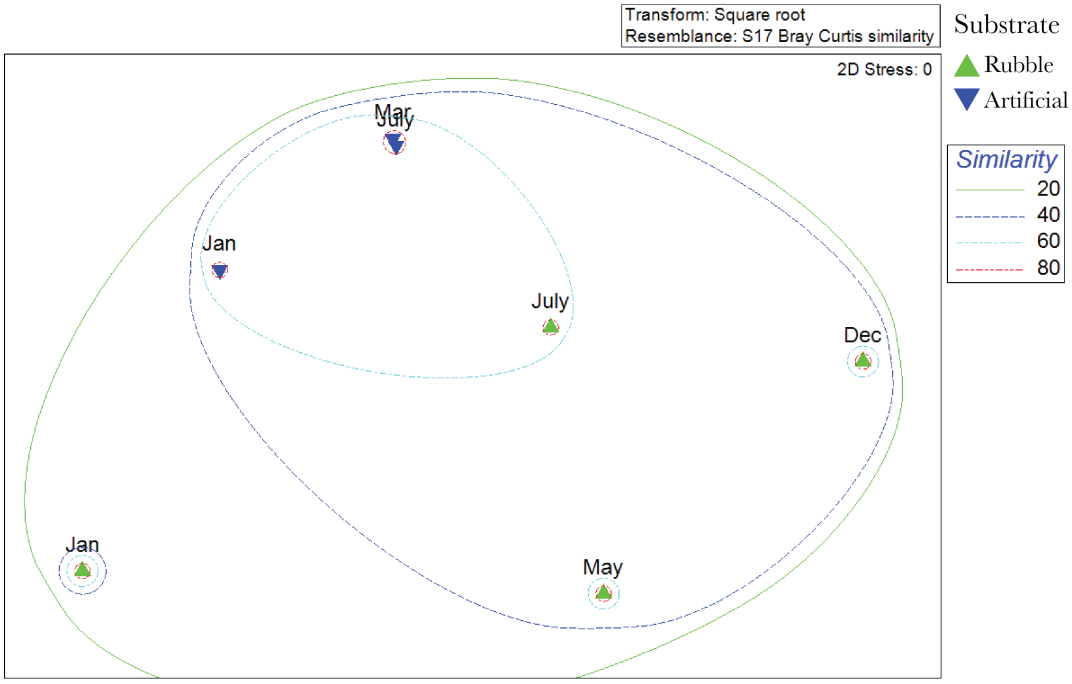
Supplemental Figures 4.6 Plots of *Gambierdiscus* Community Composition by Site and Month (Site 1 and Site 4)

MDS plots of *Gambierdiscus* community composition by site from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid.

Site 2 Bottom Lagoon



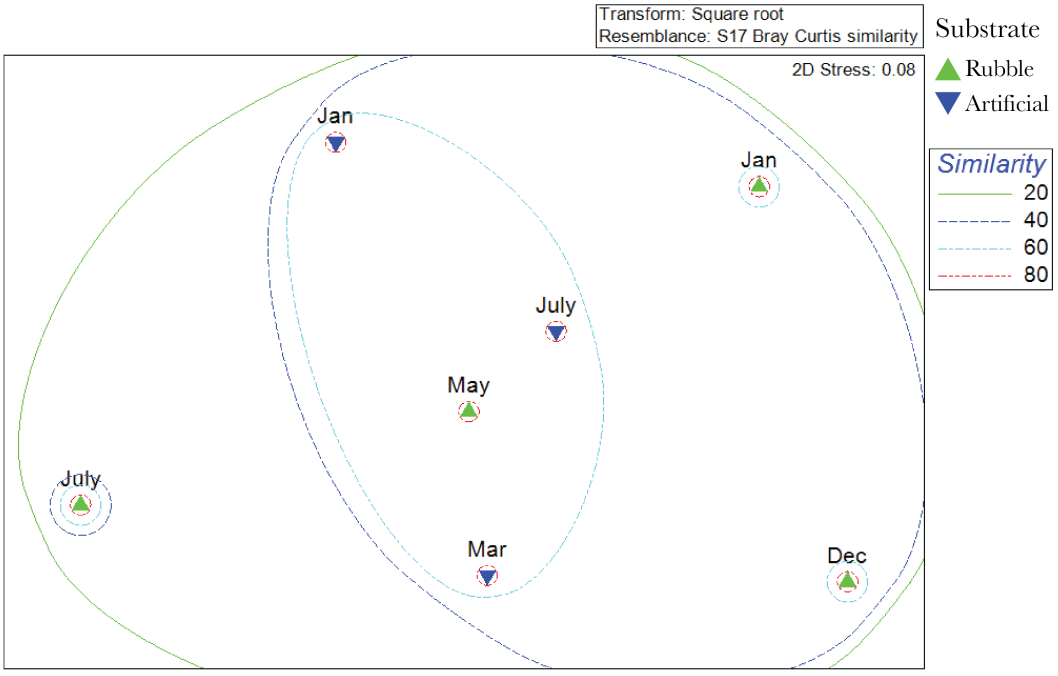
Site 5 Control 1



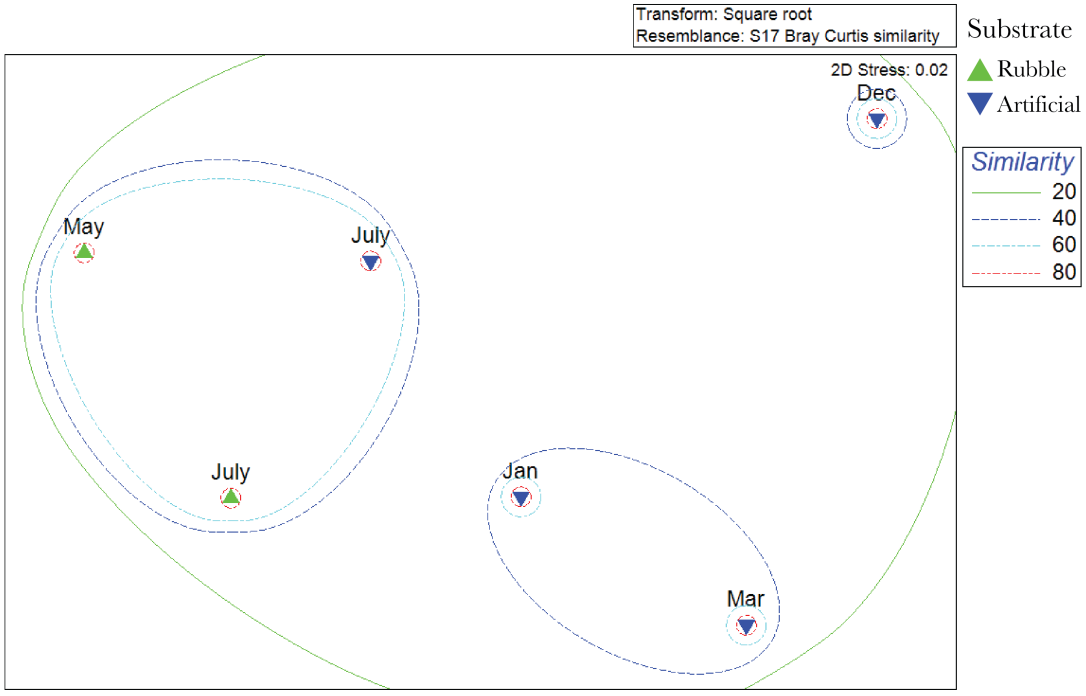
Supplemental Figures 4.7 Plots of *Gambierdiscus* Community Composition by Site and Month (Site 2 and Site 5)

MDS plots of *Gambierdiscus* community composition by site from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid.

Site 3 Control 2



Site 6 Hot Pool



Supplemental Figures 4.8 Plots of *Gambierdiscus* Community Composition by Site and Month (Site 3 and Site 6)

MDS plots of *Gambierdiscus* community composition by site from both rubble and artificial substrate data. Hierarchical clustering analysis from a Bray-Curtis similarity index resemblance matrix was overlaid.

References

Anderson, D. M. (1995). Identification of Harmful Algal Species Using Molecular Probes: An Emerging Perspective. *Harmful Marine Algal Blooms*.

Anderson, D. M., Alpermann, T. J., Cembella, A. D., Collos, Y., Masseret, E., & Montresor, M. (2012). The globally distributed genus Alexandrium: multifaceted roles in marine ecosystems and impacts on human health. *Harmful Algae*, *14*, 10–35. doi:10.1016/j.hal.2011.10.012

Anderson, D. M., Kulis, D. M., Keafer, B. a., Gribble, K. E., Marin, R., & Scholin, C. a. (2005). Identification and enumeration of Alexandrium spp. from the Gulf of Maine using molecular probes. *Deep Sea Research Part II: Topical Studies in Oceanography*, *52*(19-21), 2467–2490. doi:10.1016/j.dsr2.2005.06.015

Andersson, J. O. (2009). Gene transfer and diversification of microbial eukaryotes. *Annu Rev Microbiol*, *63*, 177–193. doi:10.1146/annurev.micro.091208.073203

Archibald, J. M. (2009). The Puzzle of Plastid Evolution. *Current Biology*, *19*(2), R81–R88. doi:10.1016/j.cub.2008.11.067

Baden, D., Fleming, L. E., & Bean, J. A. (1995). Marine toxins. In F. A. DeWolff (Ed.), *Handbook of Clinical Neurology: Intoxications of the Nervous System. Part II: Natural Toxins and Drugs.* (pp. 141–175). Amsterdam, Netherlands: Elsevier Press.

Bagnis, R., Kuberski, T., & Laugier, S. (1979). Clinical observations on 3,009 cases of ciguatera (fish poisoning) in the South Pacific. *The American Journal of Tropical Medicine and Hygiene*, *28*(6), 1067–73. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/574366

Barshis, D. J., Ladner, J. T., Oliver, T. A., & Palumbi, S. R. (2014). Lineage-Specific Transcriptional Profiles of Symbiodinium spp. Unaltered by Heat Stress in a Coral Host. *Molecular Biology and Evolution*, *31*(6), 1343–1352. doi:10.1093/molbev/msu107

Bentlage, B., Rogers, T. S., Bachvaroff, T. R., & Delwiche, C. F. (2016). Complex Ancestries of Isoprenoid Synthesis in Dinoflagellates. *Journal of Eukaryotic Microbiology*, *63*(1), 123–137. doi:10.1111/jeu.12261

Catterall, W. A. (2012). Voltage-gated sodium channels at 60: structure, function and pathophysiology. *The Journal of Physiology*, *590*(11), 2577–89. doi:10.1113/jphysiol.2011.224204

Cembella, A. D. (1998). Ecophysiology and metabolism of paralytic shellfish toxins in marine microalgae. In *Physiological Ecology of Harmful Algal Blooms. NATO-Advanced Study Institute Series* (V. 41., pp. 381–404). Heidelberg: Springer-Verlag.

Chinain, M., Faust, M. A., & Pauillac, S. (1999). Morphology And Molecular Analyses of Three Toxic Species of Gambierdiscus (Dinophyceae): G. pacificus. sp. nov., G. australes, sp. nov., and G. polynesiensis, sp. nov. *Journal of Phycology*, *35*, 1282–1296.

Clarke, K. R. (1993). Non-parametric multivariate analyses of changes in community structure. *Austral Ecology*, *18*(1), 117–143. doi:10.1111/j.1442-9993.1993.tb00438.x

Cooper, J. T., Sinclair, G., & Wawrik, B. (2016). Transcriptome Analysis of Scrippsiella trochoidea CCMP 3099 Reveals Physiological Changes Related to Nitrate Depletion. *Frontiers in Microbiology*, *7*, 639. doi:10.3389/fmicb.2016.00639

Dickey, R. W., & Plakas, S. M. (2010). Ciguatera: a public health perspective. *Toxicon : Official Journal of the International Society on Toxinology*, *56*(2), 123–36. doi:10.1016/j.toxicon.2009.09.008

Dorrell, R. G., & Howe, C. J. (2015). Integration of plastids with their hosts: Lessons learned from dinoflagellates. *Proceedings of the National Academy of Sciences*, *112*(33), 10247–10254. doi:10.1073/pnas.1421380112

EFSA, E. F. S. A. (2010). *Scientific opinion on marine biotoxins in shellfish – emerging toxins: ciguatoxin group*. *EFSA J.* (Vol. 8).

Fleming, L. E., Baden, D. G., Bean, J. A., Weisman, R., & Blythe, D. G. (1998). Marine Seafood Toxin Diseases: Issues In Epidemiology & Community Outreach. In B. Reguera, J. Blanco, M. L. Fernandez, & T. Wyatt (Eds.), *Harmful Algae* (pp. 245–248). Paris: Xunta de Galicia and Intergovernmental Oceanographic Commission of UNESCO. Retrieved from http://yyy.rsmas.miami.edu/groups/niehs/mfbsc/science/pdf/SeafoodToxinDiseasesIssues.pdf

Fleming, L. E., Broad, K., Clement, A., Dewailly, E., Elmir, S., Knap, A., … Walsh, P. (2006). Oceans and human health: Emerging public health risks in the marine environment. *Marine Pollution Bulletin*, *53*(10), 545–560. doi:10.1016/j.marpolbul.2006.08.012

Fraga, S., & Rodríguez, F. (2014). Genus Gambierdiscus in the Canary Islands (NE Atlantic Ocean) with Description of Gambierdiscus silvae sp . nov ., a New Potentially Toxic Epiphytic Benthic Dinoflagellate. *Annals of Anatomy*, *165*(6), 839–853. doi:10.1016/j.protis.2014.09.003

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–52. doi:10.1038/nbt.1883

Guillard, R. R. L. (1973). Division rates. In J. Stein (Ed.), *Handbook of phycological methods: culture methods and growth measurements.* (pp. 289–311). Cambridge: Cambridge University Press.

Hackett, J. D., Wisecaver, J. H., Brosnahan, M. L., Kulis, D. M., Anderson, D. M., Bhattacharya, D., … Erdner, D. L. (2012). Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. *Molecular Biology and Evolution*, *30*(1), 70–8. doi:10.1093/molbev/mss142

Hales, S., Weinstein, P., & Woodward, A. (1999). Ciguatera (Fish Poisoning), El Niño , and Pacific Sea Surface Temperatures. *Ecosystem Health*, *5*(1), 20–25.

Hamilton, B., Hurbungs, M., Jones, A., & Lewis, R. J. (2002). Multiple ciguatoxins present in Indian Ocean reef fish. *Toxicon*, *40*(9), 1347–1353. doi:10.1016/S0041-0101(02)00146-0

Holland, W. C., Litaker, R. W., Tomas, C. R., Kibler, S. R., Place, A. R., Davenport, E. D., & Tester, P. a. (2013). Differences in the toxicity of six Gambierdiscus (Dinophyceae) species measured using an in vitro human erythrocyte lysis assay. *Toxicon : Official Journal of the International Society on Toxinology*, *65*(January), 15–33. doi:10.1016/j.toxicon.2012.12.016

Jacobson, D. M., & Anderson, D. M. (1996). WIDESPREAD PHAGOCYTOSIS OF CILIATES AND OTHER PROTISTS BY MARINE MIXOTROPHIC AND HETEROTROPHIC THECATE DINOFLAGELLATES1. *Journal of Phycology*, *32*(2), 279–285. doi:10.1111/j.0022-3646.1996.00279.x

Jennings, S., & Kaiser, M. J. (1998). The Effects of Fishing on Marine Ecosystems. In *Advances in Marine Biology* (Vol. 34, pp. 201–352). doi:10.1016/S0065-2881(08)60212-6

Jensen, S.-K., Lacaze, J.-P., Hermann, G., Kershaw, J., Brownlow, A., Turner, A., & Hall, A. (2015). Detection and effects of harmful algal toxins in Scottish harbour seals and potential links to population decline. *Toxicon*, *97*, 1–14. doi:10.1016/j.toxicon.2015.02.002

Jeong, H. J., Yoo, Y. Du, Kim, J. S., Seong, K. A., Kang, N. S., & Kim, T. H. (2010). Growth, feeding and ecological roles of the mixotrophic and heterotrophic dinoflagellates in marine planktonic food webs. *Ocean Science Journal*, *45*(2), 65–91. doi:10.1007/s12601-010-0007-2

John, U., Litaker, R. W., Montresor, M., Murray, S., Brosnahan, M. L., & Anderson, D. M. (2014). Formal revision of the alexandrium tamarense species complex (dinophyceae) taxonomy: The introduction of five species with emphasis on molecular-based (rDNA) classification. *Protist*, *165*(6), 779–804. doi:10.1016/j.protis.2014.10.001

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., & Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation, *44*(October 2015), 457–462. doi:10.1093/nar/gkv1070

Kao, C. Y., & Nishiyama, A. (1965). Actions of saxitoxin on peripheral neuromuscular systems. *The Journal of Physiology*, *180*(1), 50–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5860225

Katz, L. A., Juhas, M., Palmer, K., Kos, V., Gilmore, M., Le, P., … Bhattacharya, D. (2015). Recent events dominate interdomain lateral gene transfers between prokaryotes and eukaryotes and, with the exception of endosymbiotic gene transfers, few ancient transfer events persist. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *370*(1678), 20140324. doi:10.1098/rstb.2014.0324

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., … Drummond, A. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics (Oxford, England)*, *28*(12), 1647–9. doi:10.1093/bioinformatics/bts199

Keeling, P. J., & Palmer, J. D. (2008). Horizontal gene transfer in eukaryotic evolution. *Nature Reviews. Genetics*, *9*(8), 605–18. doi:10.1038/nrg2386

Keeling, P. J., & Slamovits, C. H. (2005). Causes and effects of nuclear genome reduction. *Current Opinion in Genetics & Development*, *15*(6), 601–608. doi:10.1016/j.gde.2005.09.003

Kellmann, R., Mihali, T. K., Young, J. J., Pickford, R., Pomati, F., Neilan, B. A., … Neilan, B. A. (2008). Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Applied and Environmental Microbiology*, *74*(13), 4044–53. doi:10.1128/AEM.00353-08

Kellmann, R., Stüken, A., Orr, R. J. S., Svendsen, H. M., & Jakobsen, K. S. (2010). Biosynthesis and molecular genetics of polyketides in marine dinoflagellates. *Marine Drugs*, *8*(4), 1011–48. doi:10.3390/md8041011

Kibler, S. R., Litaker, W. R., Holland, W. C., Vandersea, M. W., & Tester, P. A. (2012). Growth of eight Gambierdiscus (Dinophyceae) species: Effects of temperature, salinity and irradiance. *Harmful Algae*, *19*(SEPTEMBER 2012), 1–14. doi:10.1016/j.hal.2012.04.007

Kibler, S. R., Tester, P. A., Kunkel, K. E., Moore, S. K., & Litaker, R. W. (2015). Effects of ocean warming on growth and distribution of dinoflagellates associated with ciguatera fish poisoning in the Caribbean. *Ecological Modelling*, *316*, 194–210. doi:10.1016/j.ecolmodel.2015.08.020

Kohlerl, S. T., & Kohler, C. (1992). Dead bleached coral provides new surfaces for dinoflagellates implicated in ciguatera fish poisonings. *Environmental Biology of Fishes*, *35*, 413–416.

Ku, C., Nelson-Sathi, S., Roettger, M., Sousa, F. L., Lockhart, P. J., Bryant, D., … Martin, W. F. (2015). Endosymbiotic origin and differential loss of eukaryotic genes. *Nature*, *524*(7566), 427–432. doi:10.1038/nature14963

Lange, W. R., Lipkin, K. M., & Yang, G. C. (2008). Can ciguatera be a sexually transmitted disease? *http://dx.doi.org/10.3109/15563658909038583*.

Lasley-rasher, R. S., Nagel, K., Angra, A., & Yen, J. (2016). Intoxicated copepods : ingesting toxic phytoplankton leads to risky behaviour.

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., & Prohaska, S. J. (2011). Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics*, *12*, 124. doi:10.1186/1471-2105-12-124

Lehane, L., & Lewis, R. J. (2000). Ciguatera: recent advances but the risk remains. *International Journal of Food Microbiology*, *61*(2-3), 91–125. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11078162

Lewis, R. J., & Holmes, M. J. (1993). Origin and transfer of toxins involved in ciguatera. *Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology*, *106*(3), 615–628. doi:10.1016/0742-8413(93)90217-9

Lewis, R. J., Vernoux, J.-P., & Brereton, I. M. (1998). Structure of Caribbean ciguatoxin isolated from Caranx latus. *Journal of the American Chemical Society*, *120*(24), 5914–5920. doi:10.1021/ja980389e

Li, W., & Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, *22*(13), 1658–9. doi:10.1093/bioinformatics/btl158

Lilly, E. L., Halanych, K. M., & Anderson, D. M. (2005). Phylogeny, biogeography, and species boundaries within the Alexandrium minutum group. *Harmful Algae*, *4*(6), 1004–1020. doi:10.1016/j.hal.2005.02.001

Lilly, E. L., Halanych, K. M., & Anderson, D. M. (2007). Species boundaries and global biogeography of the Alexandrium tamarense complex (Dinophyceae) 1. *Journal of Phycology*, *43*(6), 1329–1338. doi:10.1111/j.1529-8817.2007.00420.x

Litaker, R. W., Vandersea, M. W., Faust, M. a., Kibler, S. R., Chinain, M., Holmes, M. J., … Tester, P. a. (2009). Taxonomy of Gambierdiscus including four new species, Gambierdiscus caribaeus, Gambierdiscus carolinianus, Gambierdiscus carpenteri and Gambierdiscus ruetzleri (Gonyaulacales, Dinophyceae). *Phycologia*, *48*(5), 344–390. doi:10.2216/07-15.1

Llewellyn, L. E. (2006). Saxitoxin, a toxic marine natural product that targets a multitude of receptors. *Natural Product Reports*, *23*(2), 200–222. doi:10.1039/b501296c

Llewellyn, L. E. (2010). Revisiting the association between sea surface temperature and the epidemiology of fish poisoning in the South Pacific: reassessing the link between ciguatera and climate change. *Toxicon : Official Journal of the International Society on Toxinology*, *56*(5), 691–7. doi:10.1016/j.toxicon.2009.08.011

Lobel, P. S., Anderson, D. M., & Durand-Clement, M. (1988). Assessment of Ciguatera Dinoflagellate Populations: Sample Variability and Algal Substrate Selection. *Biological Bulletin*, *175*(1), 94–101.

Méheust, R., Zelzion, E., Bhattacharya, D., Lopez, P., & Bapteste, E. (2016). Protein networks identify novel symbiogenetic genes resulting from plastid endosymbiosis, *113*(13). doi:10.1073/pnas.1517551113

Mihali, T. K., Carmichael, W. W., & Neilan, B. A. (2011). A Putative Gene Cluster from a Lyngbya wollei Bloom that Encodes Paralytic Shellfish Toxin Biosynthesis, *6*(2). doi:10.1371/journal.pone.0014657

Mihali, T. K., Kellmann, R., & Neilan, B. a. (2009). Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in Anabaena circinalis AWQC131C and Aphanizomenon sp. NH-5. *BMC Biochemistry*, *10*, 8. doi:10.1186/1471-2091-10-8

Min, X. J., Butler, G., Storms, R., & Tsang, A. (2005). OrfPredictor : predicting protein-coding regions in EST-derived sequences, *33*, 677–680. doi:10.1093/nar/gki394

Moustafa, A., Loram, J. E., Hackett, J. D., Anderson, D. M., Plumley, F. G., Bhattacharya, D., & Gerald, F. (2009). Origin of saxitoxin biosynthetic genes in cyanobacteria. *PLoS ONE*, *4*(6). doi:10.1371/journal.pone.0005758

Murata, M., Legrand, A. M., Ishibashi, Y., Fukui, M., & Yasumoto, T. (1990). Structures and configurations of ciguatoxin from the moray eel Gymnothorax javanicus and its likely precursor from the dinoflagellate Gambierdiscus toxicus. *Journal of the American Chemical Society*, *112*(11), 4380–4386. doi:10.1021/ja00167a040

Murray, S. A., Diwan, R., Orr, R. J. S., Kohli, G. S., & John, U. (2015). Gene duplication, loss and selection in the evolution of saxitoxin biosynthesis in alveolates. *Molecular Phylogenetics and Evolution*, *92*, 165–180. doi:10.1016/j.ympev.2015.06.017

Murray, S. A., Mihali, T. K., & Neilan, B. A. (2011). Extraordinary conservation, gene loss, and positive selection in the evolution of an ancient neurotoxin. *Molecular Biology and Evolution*, *28*(3), 1173–1182. doi:10.1093/molbev/msq295

Murray, S. A., Wiese, M., Stüken, A., Brett, S., Kellmann, R., Hallegraeff, G., … Neilan, B. A. (2011). sxtA-based quantitative molecular assay to identify saxitoxin-producing harmful algal blooms in marine waters. *Applied and Environmental Microbiology*, *77*(19), 7050–7. doi:10.1128/AEM.05308-11

Nishimura, T., Sato, S., Tawong, W., Sakanari, H., Uehara, K., Shah, M. M. R., … Adachi, M. (2013). Genetic diversity and distribution of the ciguatera-causing dinoflagellate Gambierdiscus spp. (Dinophyceae) in coastal areas of Japan. *PloS One*, *8*(4), e60882. doi:10.1371/journal.pone.0060882

Orr, R. J. S., Stüken, A., Murray, S. A., & Jakobsen, K. S. (2013a). Evolution and distribution of saxitoxin biosynthesis in dinoflagellates. *Marine Drugs*, *11*(8), 2814–28. doi:10.3390/md11082814

Orr, R. J. S., Stüken, A., Murray, S. A., & Jakobsen, K. S. (2013b). Evolutionary acquisition and loss of saxitoxin biosynthesis in dinoflagellates: The second "core" gene, sxtG. *Applied and Environmental Microbiology*, *79*(7), 2128–36. doi:10.1128/AEM.03279-12

Orr, R. J. S., Stüken, A., Rundberget, T., Eikrem, W., & Jakobsen, K. S. (2011). Improved phylogenetic resolution of toxic and non-toxic Alexandrium strains using a concatenated rDNA approach. *Harmful Algae*, *10*(6), 676–688. doi:10.1016/j.hal.2011.05.003

Park, D. L. (1994). Evolution of methods for assessing ciguatera toxins in fish. *Reviews of Environmental Contamination and Toxicology*, *136*, 1–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8029489

Parsons, M. L., Settlemier, C. J., & Bienfang, P. K. (2010). A simple model capable of simulating the population dynamics of Gambierdiscus, the benthic dinoflagellate responsible for ciguatera fish poisoning. *Harmful Algae*, *10*(1), 71–80. doi:10.1016/j.hal.2010.07.002

Rains, L. K., & Parsons, M. L. (2015). Gambierdiscus species exhibit different epiphytic behaviors toward a variety of macroalgal hosts. *Harmful Algae*, *49*, 29–39. doi:http://dx.doi.org/10.1016/j.hal.2015.08.005

Raymond, J. A., Kim, H. J., Thomas, D., Dieckmann, G., Horner, R., Kropuenske, L., … Lovley, D. (2012). Possible Role of Horizontal Gene Transfer in the Colonization of Sea Ice by Algae. *PLoS ONE*, *7*(5), e35968. doi:10.1371/journal.pone.0035968

Richlen, M. L., Morton, S. L., Barber, P. H., & Lobel, P. S. (2008). Phylogeography, morphological variation and taxonomy of the toxic dinoflagellate Gambierdiscus toxicus (Dinophyceae). *Harmful Algae*, *7*(5), 614–629. doi:10.1016/j.hal.2007.12.020

Robertson, G., Schein, J., Chiu, R., Corbett, R., Field, M., Jackman, S. D., … Birol, I. (2010). De novo assembly and analysis of RNA-seq data. *Nature Methods*, *7*(11), 909–12. doi:10.1038/nmeth.1517

Rongo, T., Bush, M., & Van Woesik, R. (2009). Did ciguatera prompt the late Holocene Polynesian voyages of discovery? *Journal of Biogeography*, *36*(8), 1423–1432. doi:10.1111/j.1365-2699.2009.02139.x

Ruff, T. A. (1989). Ciguatera in the Pacific: A Link with Military Activities. *The Lancet*, 201–205.

Scheuer, P. J., Takahashi, W., Tsutsumi, J., & Yoshida, T. (1967). Ciguatoxin: isolation and chemical nature. *Science (New York, N.Y.)*, *155*(3767), 1267–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6018649

Scholin, C. A., & Anderson, D. M. (1993). *Population analysis of toxic and nontoxic Alexandrium species using ribosomal RNA signature sequences*. (T. J. Smayda & Y. Shimizu, Eds.)*Toxic Phytoplankton Blooms in the Sea*. Amsterdam: Elsevier.

Scholin, C. A., Herzog, M., Sogin, M., & Anderson, D. M. (1994). Identification of group- and strain-specific genetic markers for globally distributed Alexandrium (Dinophyceae). II. Sequence analysis of a fragment of the LSU rRNA gene. *Journal of Phycology*, *30*(6), 999–1011. doi:10.1111/j.0022-3646.1994.00999.x

Schönknecht, G., Weber, A. P. M., & Lercher, M. J. (2014). Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. *BioEssays*, *36*(1), 9–20. doi:10.1002/bies.201300095

Schulz, M. H., Zerbino, D. R., Vingron, M., & Birney, E. (2012). Oases : robust de novo RNA-seq assembly across the dynamic range of expression levels, *28*(8), 1086–1092. doi:10.1093/bioinformatics/bts094

Selander, E., Kubanek, J., Hamberg, M., Andersson, M. X., Cervin, G., & Pavia, H. (2015). Predator lipids induce paralytic shellfish toxins in bloom-forming algae. *Proc Natl Acad Sci U S A*, *112*(20), 6395–6400. doi:10.1073/pnas.1420154112

Shimizu, Y., Microalgae, M., Derivatives, S., Biolumlnescenca, C. I., Cycles, C., & Metabolites, D. (1993). Microalgal Metabolites. *Chem. Rev.*, *93*, 1685–1698.

Shimizu, Y., Norte, M., Hori, A., Genenah, A., & Kobayashi, M. (1984). Biosynthesis of Saxitoxin Analogues: The Unexpected Pathway. *Journal of American Chemical Society*, *106*, 6433–6434.

Shumway, S. E., Allen, S. M., & Dee Boersma, P. (2003). Marine birds and harmful algal blooms: sporadic victims or under-reported events? *Harmful Algae*, *2*(1), 1–17. doi:10.1016/S1568-9883(03)00002-7

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V, & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)*, *31*(19), 3210–2. doi:10.1093/bioinformatics/btv351

Slamovits, C. H., & Keeling, P. J. (2008). Widespread recycling of processed cDNAs in dinoflagellates. *Current Biology*, *18*(13), 550–552. doi:10.1016/j.cub.2008.04.054

Soucy, S. M., Huang, J., & Gogarten, J. P. (2015). Horizontal gene transfer: building the web of life. *Nature Reviews. Genetics*, *16*(8), 472–82. doi:10.1038/nrg3962

Stamatakis, A. (2014). RAxML Version 8 : A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies, 2010–2011.

Stucken, K., John, U., Cembella, A., Murillo, A. A., Soto-liebe, K., & Juan, J. (2010). The Smallest Known Genomes of Multicellular and Toxic Cyanobacteria : Comparison , Minimal Gene Sets for Linked Traits and the Evolutionary Implications, *5*(2). doi:10.1371/journal.pone.0009235

Stüken, A., Dittami, S. M., Eikrem, W., McNamee, S., Campbell, K., Jakobsen, K. S., & Edvardsen, B. (2013). Novel hydrolysis-probe based qPCR assay to detect saxitoxin transcripts of dinoflagellates in environmental samples. *Harmful Algae*, *28*, 108–117. doi:10.1016/j.hal.2013.06.003

Stüken, A., Orr, R. J. S., Kellmann, R., Murray, S. A., Neilan, B. A., Jakobsen, K. S., … Jakobsen, K. S. (2011). Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. *PloS One*, *6*(5), e20096. doi:10.1371/journal.pone.0020096

Suikkanen, S., Kremp, A., Hautala, H., & Krock, B. (2013). Paralytic shellfish toxins or spirolides? The role of environmental and genetic factors in toxin production of the Alexandrium ostenfeldii complex. *Harmful Algae*, *26*, 52–59. doi:10.1016/j.hal.2013.04.001

Tester, P. A., Feldman, R. L., Nau, A. W., Kibler, S. R., & Wayne Litaker, R. (2010). Ciguatera fish poisoning and sea surface temperatures in the Caribbean Sea and the West Indies. *Toxicon*, *56*(5), 698–710.

Tester, P. A., Kibler, S. R., Holland, W. C., Usup, G., Vandersea, M. W., Leaw, C. P., … Litaker, R. W. (2014). Sampling harmful benthic dinoflagellates: Comparison of artificial and natural substrate methods. *Harmful Algae*, *39*, 8–25. doi:10.1016/j.hal.2014.06.009

Tester, P. a., Vandersea, M. W., Buckel, C. a., Kibler, S. R., Holland, W. C., Davenport, E. D., … Litaker, R. W. (2013). Gambierdiscus (Dinophyceae) species diversity in the Flower Garden Banks National Marine Sanctuary, Northern Gulf of Mexico, USA. *Harmful Algae*, *29*, 1–9. doi:10.1016/j.hal.2013.07.001

Tosteson, T. . (2004). Caribbean ciguatera: a changing paradigm. *Revista de Biología Tropical*, *52*, 109–113.

Tsuchiya, S., Cho, Y., Konoki, K., Nagasawa, K., Oshima, Y., & Yotsu-Yamashita, M. (2016). Biosynthetic route towards saxitoxin and shunt pathway. *Scientific Reports*, *6*(October 2015), 20340. doi:10.1038/srep20340

Vandersea, M. W., Kibler, S. R., Holland, W. C., Tester, P. a., Schultz, T. F., Faust, M. a., … Wayne Litaker, R. (2012). Development of Semi-Quantitative Pcr Assays for the Detection and Enumeration of Gambierdiscus Species (Gonyaulacales, Dinophyceae)1. *Journal of Phycology*, *48*(4), no–no. doi:10.1111/j.1529-8817.2012.01146.x

Villareal, T. A., Hanson, S., Qualia, S., Jester, E. L. E., Granade, H. R., & Dickey, R. W. (2007). Petroleum production platforms as sites for the expansion of ciguatera in the northwestern Gulf of Mexico, *6*, 253–259. doi:10.1016/j.hal.2006.08.008

Watras, C. J., Chisholm, S. W., & Anderson, D. M. (1982). Regulation of growth in an estuarine clone of Gonyaulax tam arensis Lebour: Salinity-dependent temperature responses. *Journal of Experimental Marine Biology and Ecology*, *62*(1), 25–37. doi:10.1016/0022-0981(82)90214-3

Wiegner, T. N., Mokiao-Lee, A. U., & Johnson, E. E. (2016). Identifying nitrogen sources to thermal tide pools in Kapoho, Hawai'i, U.S.A, using a multi-stable isotope approach. *Marine Pollution Bulletin*, *103*(1-2), 63–71. doi:10.1016/j.marpolbul.2015.12.046

Wiese, M., Murray, S. A., Alvin, A., & Neilan, B. A. (2014). Gene expression and molecular evolution of sxtA4 in a saxitoxin producing dinoflagellate Alexandrium catenella. *Toxicon : Official Journal of the International Society on Toxinology*, *92*(October), 102–112. doi:10.1016/j.toxicon.2014.09.015

Wisecaver, J. H., Brosnahan, M. L., & Hackett, J. D. (2013). Horizontal gene transfer is a significant driver of gene innovation in dinoflagellates. *Genome Biology and Evolution*, *5*(12), 2368–81. doi:10.1093/gbe/evt179

Wisecaver, J. H., & Hackett, J. D. (2011). Dinoflagellate genome evolution. *Annual Review of Microbiology*, *65*, 369–87. doi:10.1146/annurev-micro-090110-102841

Xu, Y., Richlen, M. L., Liefer, J. D., Robertson, A., & Kulis, D. (2016). Influence of Environmental Variables on Gambierdiscus spp . (Dinophyceae) Growth and Distribution, 1–30. doi:10.1371/journal.pone.0153197

Xu, Y., Richlen, M. L., Morton, S. L., Mak, Y. L., Chan, L. L., Tekiau, A., & Anderson, D. M. (2014). Distribution, abundance and diversity of Gambierdiscus spp. from a ciguatera-endemic area in Marakei, Republic of Kiribati. *Harmful Algae*, *34*, 56–68. doi:10.1016/j.hal.2014.02.007

Yasumoto, T., Nakajima, I., Bagnis, R., & Adachi, R. (1977). Finding of a dinoflagellate as a likely culprit of ciguatera. *Bull. Jap. Soc. Sci. Fish*, *43*, 1021–1026.

Zdobnov, E. M., & Apweiler, R. (2001). InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics (Oxford, England)*, *17*(9), 847–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11590104

Zheng, D., Alm, E. W., Stahl, D. A., & Raskin, L. (1996). Characterization of universal small-subunit rRNA hybridization probes for quantitative molecular microbial ecology studies. *Applied and Environmental Microbiology*, *62*(12), 4504–4513. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8953722

Zimmer, R. K., & Ferrer, R. P. (2007). Neuroecology, chemical defense, and the keystone species concept. *The Biological Bulletin*, *213*(3), 208–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18083963